

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2006, 10:53:35 ; Search time 14 Seconds  
(without alignments)  
2219.858 Million cell updates/sec

Title: US-10-634-905-2  
Perfect score: 1705  
Sequence: 1 MKSALTRFFILLPWILVI.....EKNVLANEPKYLDTVKIEV 323

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_80.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1526.5	89.5	324	JG0163	glucuronyltransfer
2	777.5	45.6	347	JC7828	glucuronyltransfer
3	529.5	31.1	356	T27733	hypothetical prote
4	365	21.4	259	T20205	hypothetical prote
5	361	21.2	290	T24926	hypothetical prote
6	355	20.8	248	T20027	hypothetical prote
7	347.5	20.4	226	T20447	hypothetical prote
8	336	19.7	325	T24737	hypothetical prote
9	180	10.6	351	D84788	hypothetical prote
10	169	9.9	544	T85435	UDP-glucuronyltran
11	111	6.5	666	T70803	hypothetical prote
12	107	6.3	516	T49422	RAD51 related prot
13	102	6.0	786	A47547	serine proteinase
14	101	5.9	92	T29701	hypothetical prote
15	99.5	5.8	813	S33261	protein kinase lin
16	97	5.7	586	T29695	hypothetical prote
17	96	5.6	767	S41479	DNA-binding protei
18	96	5.6	2282	T42717	DNA-binding protei
19	94	5.5	1360	T32833	hypothetical prote
20	92	5.4	497	S22708	homeotic protein e
21	92	5.4	1396	T24453	translation initia
22	91.5	5.4	367	H83088	membrane-bound lyt
23	91.5	5.4	445	A75376	probable oligosacc
24	91	5.3	1466	T32422	hypothetical prote
25	90.5	5.3	285	S08491	hypothetical prote
26	90.5	5.3	520	F0LJGL	gag polyprotein -
27	90.5	5.3	1520	T00723	hypothetical prote
28	90.5	5.3	4135	T42629	tenascin-X - bovin
29	89.5	5.2	366	AG2060	hypothetical prote

30	89.5	5.2	497	2	F83634	hypothetical prote
31	89.5	5.2	591	2	G96734	spore coat protein
32	89	5.2	428	2	A83005	conserved hypothet
33	88.5	5.2	448	2	D87146	conserved hypothet
34	88.5	5.2	818	2	A59433	KIAA0672 protein (
35	88	5.2	727	2	AD1868	hypothetical prote
36	87.5	5.1	364	2	C87455	alanine racemase (
37	87	5.1	1258	2	JC5765	inositol polyphosp
38	86	5.0	964	2	D59404	plectin isoform p1
39	85.5	5.0	370	2	AG0359	probable membrane-
40	85.5	5.0	1091	2	S33596	protein-tyrosine k
41	85.5	5.0	1892	2	T18314	hypothetical prote
42	85.5	5.0	4687	1	A39638	plectin - rat
43	85	5.0	1377	2	C65159	rhaA protein precu
44	84.5	5.0	147	2	S37485	gene mgi1 protein
45	84.5	5.0	264	2	AH2041	hypothetical prote

ALIGNMENTS

RESULT 1  
JG0163

glucuronyltransferase - rat  
C/Species: Rattus norvegicus (Norway rat)  
C/Date: 23-Jul-1999 #sequence\_revision 23-Jul-1999 #text\_change 09-Jul-2004  
C/Accession: JG0163

R/Seiki, T.; Oka, S.; Terayama, K.; Imiya, K.; Kawasaki, T.

Biochem. Biophys. Res. Commun. 255, 182-187, 1999  
A/Title: Molecular cloning and expression of a second glucuronyltransferase involved in t  
A/Reference number: JG0163; MUID:99185317; PMID:10082676

A/Accession: JG0163

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-324 <SEI>

A/Cross-references: UNIPROT:Q9Z137; UNIPARC:UPI000012670D; DDBJ:AB010441; NID:94519213; I

Query Match 89.5%; Score 1526.5; DB 2; Length 324;  
Best Local Similarity 89.8%; Pred. No. 5.8e-119;  
Matches 291; Conservative 10; Mismatches 22; Indels 1; Gaps 1;

Qy	1	MKSALTRFFILLPWILVIIMLDVDRTRPVPLTPRPYFSPVAVGGRGARLPLRRGGPA	60
Db	1	MKSALCNRRFFILLPWILVIIMLDVDRPRPAPQLTSRPYFSPHTVGGGSRVPLRRSSPG	60
Qy	61	H-GTQGNQSRQPQPEPOLPTIYAITPYSPVQKAELTRLANTRQVAQLHWILVEDA	119
Db	61	RDAAEKRNESRPQLOPEPLPTIYAITPYSPVQKAELTRLANTRQVAQLHWILVEDR	120
Qy	120	AARSELVSRFLARAGLPSTHLVPTPRRYKRPGLPRATEORNAGLAWLRQHQAOPG	179
Db	121	ATRSELVSSFLARAGLPNTLHVPTPRRYKRPGLPRATEORNAGLAWLRQHQAOPG	180
Qy	180	VLFFADDDNTYSLELFPQEMRTTRKVSVPVGLVGGRRYRPLRVENGKVGWYTGWRADRP	239
Db	181	VLFFADDDNTYSLELFPQEMRTTRKVSVPVGLVGGRRYRPLRVENGKVGWYTGWRDRP	240
Qy	240	FAIDMAGFAVSLQVILSNKAVFKRRGSGQGQESDFLQITTVEELEPKANNCTKVLVW	299
Db	241	FAIDMAGFAVSLQVILSNKAVFKRRGSGQGQESDFLQITTVDSLEPKANNCTKVLVW	300
Qy	300	HYRTEKVNLANEPKYHLDTVKIEV	323
Db	301	HYRTEKVNLANEPKYHMDTVNIEV	324

RESULT 2  
JC7828

glucuronyltransferase-P, long form - mouse

C/Species: Mus musculus (house mouse)

C/Date: 03-Jun-2002 #sequence\_revision 03-Jun-2002 #text\_change 09-Jul-2004

C/Accession: JC7828

R/Yamamoto, S.; Oka, S.; Saito-Ohara, F.; Inazawa, J.; Kawasaki, T.





Qy 81 TIYAITPTYSRPVQKAELTRLANTRFQVAOLHWILVEDAAARSELVSRFLARGLPSTHL 140  
 Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :  
 Qy 40 TIIVITPTYRINRMPDITRLSNLTSHVKNLHWIIVDGVSTVPVAVRAVLRTGLSYTYM 99  
 Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :  
 Qy 141 HVPTPRYKRPGLPRATEQRNAGLAWLRQ---RHQHQAQPGVLFFADDNDNTYSLELFOE 197  
 Db :  
 Qy 100 AHKTAQGYPAKG---WVQRTMALKFIRENTSRILNTDLREGVVYFADDNSYDLRLFND 155  
 Db :  
 Qy 198 -MRTRKVSUWVPV-----GLVGRRYERPLVENGKVKVGYTGWRAADRP 239  
 Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :  
 Qy 156 FIRNVRKLGWAVGAKQAASQNTDHSFFVYLGFGAGAAVEAPKVVDKVTSTFDALWVSKRL 215  
 Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :  
 Qy 240 FAIDMAGFAVSLQVILSNPRAVFKR---RGSQPMQESDFLKQI--TTVESELEPKA----- 290  
 Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :  
 Qy 216 FAVDMAGFAVNLKWIL--RTWAVGKTCNRGD--GAPETCLEDLGPFLEIDIEPFGYEQV 272  
 Db :  
 Qy 291 -----NNTCKVLVWHTRT 303  
 Db :  
 Qy 273 SNFKVLLLEHLTKTPALLYNSIFQNNREILVWHTRT 308  
 Db :  
 RESULT 9  
 D84788  
 Hypothetical protein At2g37090 [imported] - Arabidopsis thaliana  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
 C/Accession: D84788  
 R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;  
 cuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A/Reference number: A84420; MUID:20083487; PMID:10617197  
 A/Accession: D84788  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-351 <STO>  
 A/Cross-references: UNIPROT.Q9ZQG; UNIPARC:UIP00000A6561; GB:A5002093; NID:94371292; PI:  
 C/Genetics:  
 A/Gene: At2g37090  
 A/Map position: 2

**F85435**

UDP-glucuronyltransferase-like protein [imported] - Arabidopsis thaliana  
 C|Species: Arabidopsis thaliana (mouse-ear cress)  
 C|Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
 C|Accession: F85435  
 R|anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring  
 Nature 402, 769-777, 1999  
 A|Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
 A|Reference number: A85001; MUID:20083488; PMID:10617198  
 A|Accession: F85435  
 A|Status: preliminary  
 A|Molecule type: DNA  
 A|Residues: 1-544 <STO>  
 A|Cross-references: UNIPROT:O23194; UNIPARC:UPI0000048A7D; GB:NC\_001268; NID:g7270638; PJ  
 C|Genetics:  
 A|Gene: AT4g36890  
 A|Map position: 4

RESULT 10



Matches	30;	Conservative	14;	Mismatches	22;	Indels	12;	Gaps	5;
Qy	244	MAGFAVSLQVILSNPKAVF----	KRGSQPMQESDFLKOI-TTVBELEP----	KANNCTYK	295				
Db	1	MAGFAVNLKVV-LNSDAVFTACKRGGAP----	ETCLLEDMGLEREDIEFFGYEKDKDRE	56					
Qy	296	VLVWHHTRTKVNLANEPK	313						
Db	57	ILVWHHTKSTPNIVKSNK	74						
RESULT 15									
S33261									
protein kinase lin-45 (EC 2.7.1.1) - Caenorhabditis elegans									
C;Species: Caenorhabditis elegans									
C;Date: 22-Nov-1993 #sequence_revision 01-Sep-1995 #text_change 05-Oct-2004									
C;Accession: S33261									
R;Han, M.; Golden, A.; Han, Y.; Sternberg, P.W.									
Nature 363, 133-140, 1993									
A;Title: C. elegans lin-45 raf gene participates in let-60 ras-stimulated vulval differentiation									
A;Reference number: S33261; PMID:93247635; PMID:8483497									
A;Accession: S33261									
A;Status: preliminary									
A;Molecule type: mRNA									
A;Residues: 1-813 <HAN>									
A;Cross-references: UNIPROT:Q07292; UNIPARC:UPI000016B91E; EMBL:L15347; NID:G289715; PIR									
C;Genetics:									
A;Gene: lin-45									
A;Map position: IV									
A;Note: between unc-44 and deb-1									
C;Function:									
A;Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine-phosphate									
A;Pathway: vulval differentiation									
C;Superfamily: serine/threonine-protein kinase, A/B-Raf proto-oncogene types; protein kinase C									
C;Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase									
F;171-217/Domain: protein kinase C zinc-binding repeat homology <K22>									
F;479-748/Domain: protein kinase homology <KIN>									
F;487-495/Region: protein kinase ATP-binding motif									
F;171,198,201,217/Binding site: zinc (His, Cys, Cys) #status predicted									
F;184,187,206,209/Binding site: zinc (Cys, Cys, His, Cys) #status predicted									
F;312/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted									
F;507,525,602,604/Active site: Lys, Glu, Asp, Lys #status predicted									
F;756/Binding site: phosphate (Ser) (covalent) #status predicted									
Query Match									
Best Local Similarity 21.4%; Pred. No. 3.4;									
Matches 55; Conservative 43; Mismatches 102; Indels 57; Gaps 15;									
Qy	25	VDTRRRVPP-----	LTTPYFSPYAVRGGARLPLRRGGPAHGT----	OKRQSRPQ	72				
Db	458	LDAQRPRPPQKPHHDEWILNFEI	IQYKVGSGSGFTVYR--GEFFGTVAIKKLNVDPT	515					
Qy	73	PQPEQLPTIYAITPTYSRVPQKAE	LTRANTFRQVLAHLVDEAAARSELVSRFLAR	132					
Db	516	PSQ-----	MAAFKNEVAVLKKTRHLNV---LLFGWVREPIA----	IITQCEG	558				
Qy	133	AGLPSTLHVPTPR-RYKRPGLPRATE	QORNAGLAWLRQRH-QHORAQPGVLFFADDDNTY	190					
Db	559	SSL-YRHHVQEPREVPFEMGAII	DLKQVSLGNLYLHSHKNIHRDLKTNNIFLMDDMSTV	617					
Qy	191	SLELFQEMRTTRKVSVPVCLVGR	RYERPLVNGKVVGYTCWRA-----	DRPFA-	241				
Db	618	KIGDFGLATVTKTK---	WTVN--GGQQQQQP---TGSIL-----	WMAPEVIRMQDDNPYTP	664				
Qy	242	-IDMAGFAVSLQVILSN	257						
Db	665	QSDVYSFGICMYEILSS	681						

Search completed: May 9, 2006, 10:55:41  
Job time : 16 secs

GenCore version 5.1.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2006, 10:56:24 ; Search time 40 Seconds  
(without alignments)  
776.950 Million cell updates/sec

Title: US-10-634-905-2  
Perfect score: 323  
Sequence: 1 MKSALFTRFFILLPWILVI.....EKNVLANEKYLHDTVKIEV 323

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 1

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR 80.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	14.6	324	2 JG0163	glucuronyltransfer
2	15	4.6	347	2 JC7828	glucuronyltransfer
3	9	2.8	257	2 D84033	ribonuclease PH rp
4	8	2.5	138	2 B27863	Ac hypothetical pr
5	8	2.5	144	1 YMSWCM	macromomycin precu
6	8	2.5	356	2 T27733	hypothetical prote
7	8	2.5	427	2 T03955	probable transposa
8	8	2.5	487	1 H81109	IMP dehydrogenase
9	8	2.5	487	2 F81906	IMP dehydrogenase
10	8	2.5	807	2 T02916	hypothetical prote
11	8	2.5	839	1 TQZMCA	probable transposa
12	8	2.5	3198	2 A43426	collagen alpha 2 f
13	7	2.2	64	2 AB3560	conserved cytosoli
14	7	2.2	87	2 H64521	hypothetical prote
15	7	2.2	87	2 D71984	hypothetical prote
16	7	2.2	95	2 AD3622	transcription regu
17	7	2.2	133	2 I64784	hypothetical prote
18	7	2.2	150	2 H72603	hypothetical prote
19	7	2.2	156	2 T41316	peroxisomal membra
20	7	2.2	176	2 E69519	conserved hypothet
21	7	2.2	178	2 C38487	single-stranded DN
22	7	2.2	179	1 DDECF	threonine efflux p
23	7	2.2	206	2 AH0466	amino acid ABC tra
24	7	2.2	220	2 A81159	hypothetical prote
25	7	2.2	226	2 T20447	hypothetical prote
26	7	2.2	229	1 W4WL8	E4 protein - human
27	7	2.2	237	1 S11927	licheninase (EC 3.
28	7	2.2	238	2 E81947	probable amino aci
29	7	2.2	239	2 E89830	hypothetical prote

ALIGNMENTS

RESULT 1

JG0163 glucuronyltransferase - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 23-Jul-1999 #sequence\_revision 23-Jul-1999 #text\_change 09-Jul-2004  
C:Accession: JG0163

R:Seiki, T.; Oka, S.; Terayama, K.; Imiya, K.; Kawasaki, T.

Biochem. Biophys. Res. Commun. 255, 182-187, 1999  
A:Title: Molecular cloning and expression of a second glucuronyltransferase involved in  
A:Reference number: JG0163; MUID:99185317; PMID:10082676

A:Accession: JG0163

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-324 <SEI>

A:Cross-references: UNIPROT:Q9Z137; UNIPARC:UPI000012670D; DBJ:AB010441; NID:94519213; I

Query Match 14.6%; Score 47; DB 2; Length 324;

Best Local Similarity 100.0%; Pred. No. 5.3e-40;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 176 AQPGLFFADDNTYSLELFQEMRTTKVSVMPVGLVGGRRYERPLV 222

Db 177 AQPGLFFADDNTYSLELFQEMRTTKVSVMPVGLVGGRRYERPLV 223

RESULT 2

JC7828

glucuronyltransferase-P, long form - mouse

C:Species: Mus musculus (house mouse)

C:Date: 03-Jun-2002 #sequence\_revision 03-Jun-2002 #text\_change 09-Jul-2004

C:Accession: JC7828

R:Yamamoto, S.; Oka, S.; Saito-Ohara, F.; Inazawa, J.; Kawasaki, T.

J. Biochem. 131, 337-347, 2002

A:Title: Molecular cloning and genomic analysis of mouse glucuronyltransferase involved in

A:Reference number: JC7828; PMID:11872162; MUID:21861939

A:Contents: Brain

A:Accession: JC7828

A:Molecule type: mRNA

A:Residues: 1-347 <YAM>

A:Cross-references: UNIPROT:Q8R531; UNIPARC:UPI00000888CC; DBJ:AB055781

C:Comment: This enzyme, highly conserved during phylogenesis, is a key enzyme involved in

C:Genetics:

A:Gene: GICAT-P

A:Map position: A4-9

A:Introns: 2/1; 51/1; 220/3; 319/3

Query Match 4.6%; Score 15; DB 2; Length 347;

Best Local Similarity 100.0%; Pred. No. 4.2e-07;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 86 TPTYSRPVQAELTR 100

|||||

```

Db      103 TPTYSRPVQKALTR 117

RESULT 3
D84033
ribonuclease PH rph [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: D84033
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: D84033
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-257 <STO>
A:Cross-references: UNIPROT:Q9K8D8; UNIPARC:UPI0000134454; GB:AP001517; GB:BA0000004; NID
A:Experimental source: strain C-125
C:Genetics:
C:Superfamily: tRNA nucleotidyltransferase

Query Match      2.8%; Score 9; DB 2; Length 257;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      153 LPRATEQRN 161
Db      66 LPRATEQRN 74

RESULT 4
B27863
Ac hypothetical protein ORPa - maize (fragment)
C:Species: Zea mays (maize)
C>Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 30-Sep-1993
C:Accession: B27863
R:Kunze, R.; Stochaj, U.; Laufs, J.; Starlinger, P.
EMBO J. 6, 1555-1563, 1987
A:Title: Transcription of transposable element Activator (Ac) of Zea mays L.
A:Reference number: A91069
A:Accession: B27863
A:Molecule type: mRNA
A:Residues: 1-138 <RUN>
A:Cross-references: UNIPARC:UPI000017835B
C:Superfamily: hypothetical transposase Ac9

Query Match      2.5%; Score 8; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      71 PQQPEPQ 78
Db      113 PQQPEPQ 120

RESULT 5
YMSMCM
macromycin precursor - Streptomyces "macromyceticus"
N:Alternate names: auroomycin apoprotein (AUR)
N:Contains: macromycin (MCR)
C:Species: Streptomyces "macromyceticus"
C>Date: 31-May-1979 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C:Accession: A45766; A92409; A90212; A01812; JS0338
R:Sakata, N.; Kanbe, T.; Tanabe, M.; Hayashi, H.; Hori, M.; Hotta, K.; Hamada, M.
J. Antibiot. 42, 1704-1712, 1989
A:Title: Nucleotide sequence of the macromycin apoprotein gene and its expression in S
A:Reference number: A45766; MUID:90061944; PMID:2479629
A:Accession: A45766
A:Molecule type: DNA
A:Residues: 1-144 <SAK>
A:Cross-references: UNIPROT:P01549; UNIPARC:UPI000012EB0C; GB:D90006; NID:g217017; PIDN:

```

```

A:Experimental source: strain M480-M1
R:Samy, T.S.A.; Hahn, K.S.; Modest, E.J.; Lampman, G.W.; Keutmann, H.T.; Umezawa, H.; Her
J. Biol. Chem. 258, 183-191, 1983
A:Title: Primary structure of macromycin, an antitumor antibiotic protein.
A:Reference number: A92409; MUID:83082859; PMID:6848492
A:Accession: A92409
A:Molecule type: protein
A:Residues: 33-110, 'D', 112-144 <SAM>
A:Cross-references: UNIPARC:UPI00001125AA
R:Sawyer, T.H.; Gietzow, K.; Olson, M.O.J.; Busch, H.; Prestayko, A.W.; Crooke, S.T.
Biochem. Biophys. Res. Commun. 86, 1133-1138, 1979
A:Title: Amino terminal amino acid sequence of macromycin, a protein antitumor antibiotic
A:Reference number: A90212; MUID:79165840; PMID:155453
A:Accession: A90212
A:Molecule type: protein
A:Residues: 33-47, 'E', 49-64, 66, 'ES', 69-70, 'A', 72-77, 'P' <SAW>
A:Cross-references: UNIPARC:UPI0000173628
R:Van Roey, P.; Beerman, T.A.
Proc. Natl. Acad. Sci. U.S.A. 86, 6587-6591, 1989
A:Title: Crystal structure analysis of auroomycin apoprotein (macromycin) shows import
A:Reference number: A33726; MUID:89367294; PMID:2771945
C:Contents: annotation; X-ray crystallography, 1.5 angstroms
C:Comment: The antibiotic activity (for gram-positive bacteria) and the antitumor activit
ific inhibition of DNA synthesis, with consequent arrest of mitosis, in these susceptible
C:Genetics:
A:Gene: mcmA
C:Superfamily: macromycin
C:Keywords: antibiotic; DNA binding; pigment binding
F:1-32/Domain: signal sequence #status predicted <SIG>
F:33-144/Product: macromycin #status experimental <MAT>
F:68-78,120-125/Disulfide bonds: #status experimental

Query Match      2.5%; Score 8; DB 1; Length 144;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      127 SRFLARAG 134
Db      6 SRFLARAG 13

RESULT 6
T27733
hypothetical protein ZK1307.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T27733
R:Berks, M.
submitted to the EMBL Data Library, January 1995
A:Reference number: Z20412
A:Accession: T27733
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-356 <WIL>
A:Cross-references: UNIPROT:Q09363; UNIPARC:UPI0000135EA9; EMBL:Z47358; PIDN:CAA87436.1;
A:Experimental source: clone ZK1307
C:Genetics:
A:Gene: CESP:ZK1307.5
A:Map position: 2
A:Introns: 46/3; 81/2; 247/1

Query Match      2.5%; Score 8; DB 2; Length 356;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      298 VWHTRTEK 305
Db      320 VWHTRTEK 327

RESULT 7
T03955
probable transposase - maize transposable element Ac

```

C;Species: Zea mays (maize)  
C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C;Accession: T03955  
R;Kunze, R.; Starlinger, P.  
EMBO J. 8, 3177-3185, 1989  
A;Title: The putative transposase of transposable element Ac from Zea mays L. interacts  
A;Reference number: Z15160; MUID:90059907; PMID:2555157  
A;Accession: T03955  
A;Status: preliminary;  
A;Molecule type: DNA  
A;Residues: 1-427 <KUN>  
A;Cross-references: UNIPROT:Q98AP9; UNIPARC:UPI000009D1DC; EMBL:X01380; NID:g22489; PIDN  
C;Genetics:  
A;Mobile element: transposable element Ac  
C;Superfamily: hypothetical transposase Ac9

Query Match 2.5%; Score 8; DB 2; Length 427;  
Best Local Similarity 100.0%; Pred. No. 7.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 PQQPPEPQ 78  
Db 11 PQQPPEPQ 18  
|||||

RESULT 8  
H81109  
IMP dehydrogenase (EC 1.1.1.205) NM01201 [similarity] - Neisseria meningitidis (strain M  
N;Alternate names: IMP:NAD+ oxidoreductase; IMPDH; inosine-5'-monophosphate dehydrogenase  
C;Species: Neisseria meningitidis  
C;Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 09-Jul-2004  
C;Accession: H81109  
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.  
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
Li, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masiagnani, V.; Pizza, M.  
Science 287, 1809-1815, 2000  
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve  
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
A;Reference number: A81000; MUID:20175755; PMID:10710307  
A;Accession: H81109  
A;Molecule type: DNA  
A;Residues: 1-487 <TET>  
A;Cross-references: UNIPROT:Q9JZB5; UNIPARC:UPI000000C4660; GB:AE002468; GB:AE002098; NID  
A;Experimental source: serogroup B, strain MC58  
C;Genetics:  
A;Gene: NM01201  
C;Function:  
A;Description: catalyzes the oxidation of inosine 5'-phosphate by NAD and one molecule o  
A;Pathway: purine nucleotide biosynthesis  
C;Superfamily: inosine-5'-monophosphate dehydrogenase; CBS homology; IMP dehydrogenase a  
C;Keywords: duplication; GMP biosynthesis; NAD; oxidoreductase; purine nucleotide biosyn  
F;8-74/Domain: IMP dehydrogenase amino-terminal homology <IDHN>  
F;95-144/Domain: CBS homology <CBS1>  
F;158-206/Domain: CBS homology <CBS2>  
F;207-451/Domain: IMP dehydrogenase catalytic homology <IDHC>  
F;305/Active site: Cys #status predicted

Query Match 2.5%; Score 8; DB 1; Length 487;  
Best Local Similarity 100.0%; Pred. No. 8.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 222 VENGKVVG 229  
Db 125 VENGKVVG 132  
|||||

RESULT 9  
F81906  
IMP dehydrogenase (EC 1.1.1.205) NM01372 [similarity] - Neisseria meningitidis (strain 2  
C;Species: Neisseria meningitidis  
C;Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 09-Jul-2004  
C;Accession: F81906  
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel

; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,  
Nature 404, 502-506, 2000  
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.  
A;Reference number: A81775; MUID:20222556; PMID:10761919  
A;Accession: F81906  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-487 <PAR>  
A;Cross-references: UNIPROT:Q9JUD0; UNIPARC:UPI000000C4BAA; GB:AL162755; GB:AL157959; NID:  
A;Experimental source: serogroup A, strain Z2491  
C;Genetics:  
A;Gene: guaB; NM01372  
C;Superfamily: inosine-5'-monophosphate dehydrogenase; CBS homology; IMP dehydrogenase an  
C;Keywords: GMP biosynthesis; NAD; oxidoreductase  
F;8-74/Domain: IMP dehydrogenase amino-terminal homology <IDN>  
F;305/Active site: Cys #status predicted

Query Match 2.5%; Score 8; DB 2; Length 487;  
Best Local Similarity 100.0%; Pred. No. 8.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 222 VENGKVVG 229  
Db 125 VENGKVVG 132  
|||||

RESULT 10  
T02916  
Hypothetical protein a - maize transposable element Ac  
C;Species: Zea mays (maize)  
C;Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004  
C;Accession: T02916  
R;Kunze, R.  
submitted to the EMBL Data Library, February 1988  
A;Reference number: Z14767  
A;Accession: T02916  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-807 <KUN>  
A;Cross-references: UNIPROT:P08770; UNIPARC:UPI000016E023; EMBL:X05424; NID:g22112; PIDN  
C;Genetics:  
A;Mobile element: transposable element Ac  
A;Introns: 41/1; 490/2; 756/1; 786/3  
C;Superfamily: hypothetical transposase Ac9

Query Match 2.5%; Score 8; DB 2; Length 807;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 PQQPPEPQ 78  
Db 113 PQQPPEPQ 120  
|||||

RESULT 11  
TQ2MCA  
probable transposase - maize transposon Ac9  
C;Species: Zea mays (maize)  
C;Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 09-Jul-2004  
C;Accession: A03540; S18636  
R;Pohlman, R.F.; Fedoroff, N.V.; Messing, J.  
Cell 37, 635-643, 1984  
A;Title: The nucleotide sequence of the maize controlling element activator.  
A;Reference number: A90848; MUID:84205699; PMID:6327080  
A;Accession: A03540  
A;Molecule type: DNA  
A;Residues: 1-839 <POH>  
A;Cross-references: UNIPROT:P03010; UNIPARC:UPI00001372D4; GB:X01964; NID:g168402  
R;Feldmar, S.; Kunze, R.  
EMBO J. 10, 4003-4010, 1991  
A;Title: The ORF4 protein, the putative transposase of maize transposable element Ac, has  
A;Reference number: S18636; MUID:92097517; PMID:1661668  
A;Accession: S18636

A;Status: nucleic acid sequence not shown; not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 103-184 <FEL>  
A;Cross-references: UNIPARC:UPI0000174897  
C;Comment: This protein is coded by the transposable maize controlling element "Activator" of the *Ac* transposon.  
C;Superfamily: hypothetical transposase Ac9  
C;Keywords: DNA binding; DNA replication  
F;136-145/Region: DNA binding #status experimental

Query Match 2.5%; Score 8; DB 1; Length 839;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 PQQPPEPQ 78  
DB 58 PQQPPEPQ 65  
|||||

RESULT 12  
A43426  
collagen alpha 2 fibrillar chain precursor - sea urchin (*Strongylocentrotus purpuratus*)  
C;Species: *Strongylocentrotus purpuratus* (purple urchin)  
C;Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C;Accession: A43426  
R;Exposito, J.Y.; D'Alessio, M.; Ramirez, F.  
J. Biol. Chem. 267, 17404-17408, 1992  
A;Title: Novel amino-terminal propeptide configuration in a fibrillar procollagen underg  
A;Reference number: A43426; MUID:92381062; PMID:1380962  
A;Accession: A43426  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: nucleic acid  
A;Residues: 1-3198 <EXP>  
A;Cross-references: UNIPROT:Q26639; UNIPARC:UPI0000077B9A; GB:M92041; NID:g161448; PIDN:  
F;48-106/Domain: von Willebrand factor type C repeat homology <WMC>  
F;2978-3198/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Query Match 2.5%; Score 8; DB 2; Length 3198;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 208 PVGLVGR 215  
DB 2768 PVGLVGR 2775  
|||||

RESULT 13  
AB3560  
conserved cytosolic protein [imported] - *Brucella melitensis* (strain 16M)  
C;Species: *Brucella melitensis*  
C;Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 09-Jul-2004  
C;Accession: AB3560  
R;DelVecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,  
.; Mazur, M.; Goldsman, E.; Salkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letese  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A;Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*  
A;Reference number: AD3252; PMID:11756688  
A;Accession: AB3560  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-64 <KUR>  
A;Cross-references: UNIPROT:Q8YCK5; UNIPROT:Q8FVF2; UNIPARC:UPI0000058482; GB:AE008918;  
A;Experimental source: strain 16M  
C;Genetics:  
A;Gene: BMEI0403  
A;Map position: II

Query Match 2.2%; Score 7; DB 2; Length 64;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 RSELVSR 128

DB 33 RSELVSR 39  
|||||

RESULT 14  
H64521  
hypothetical protein HP0016 - *Helicobacter pylori* (strain 26695)  
C;Species: *Helicobacter pylori*  
C;Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 09-Jul-2004  
C;Accession: H64521  
R;Tom, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.  
Nature 388, 539-547, 1997  
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
A;Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.  
A;Reference number: A64520; MUID:97394467; PMID:9252185  
A;Accession: H64521  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-87 <TOM>  
A;Cross-references: UNIPROT:Q24861; UNIPARC:UPI00000C07CC; GB:AE000524; GB:AE000511; NID:  
Query Match 2.2%; Score 7; DB 2; Length 87;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 FILLPMI 16  
DB 36 FILLPMI 42  
|||||

RESULT 15  
D71984  
hypothetical protein jhp0014 - *Helicobacter pylori* (strain J99)  
C;Species: *Helicobacter pylori*  
A;Variety: strain J99  
C;Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
C;Accession: D71984  
R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;  
; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; J  
Nature 397, 176-180, 1999  
A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric path  
A;Reference number: A71800; MUID:99120557; PMID:9923682  
A;Accession: D71984  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-87 <ARN>  
A;Cross-references: UNIPROT:Q9ZM45; UNIPARC:UPI00000D713D; GB:AE001441; GB:AE001439; NID:  
A;Experimental source: strain J99  
C;Genetics:  
A;Gene: jhp0014

Query Match 2.2%; Score 7; DB 2; Length 87;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 FILLPMI 16  
DB 36 FILLPMI 42  
|||||

Search completed: May 9, 2006, 10:59:43  
Job time : 42 secs

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2006, 10:56:10 ; Search time 59 Seconds  
(without alignments)  
3862.469 Million cell updates/sec

Title: US-10-634-905-2

Perfect score: 323

Sequence: 1 MKSALFTRFFILLFWILVI.....EKNVLANEPKYHLDTVKIEV 323

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2166443 seqs, 705528306 residues

Word size : 1

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : UniProt 05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	323	100.0	323	1 B3GA2_HUMAN	Q9np25 homo sapien
2	323	100.0	323	2 Q599K0_PANTR	Q599k0 pan troglod
3	236	73.1	329	1 B3GA2_CANFA	Q5ca26 canis famil
4	105	32.5	326	2 Q599K1_BOVIN	Q599k1 bos taurus
5	62	19.2	304	2 Q5CA57_CHICK	Q5ca57 gallus gall
6	54	16.7	324	1 B3GA2_MOUSE	P59270 mus musculu
7	54	16.7	349	2 Q5DTT7_MOUSE	Q5dtt7 mus musculu
8	47	14.6	324	1 B3GA2_RAT	Q9z137 rattus norv
9	29	9.0	331	2 Q599J8_XENTR	Q599j8 xenopus tro
10	27	8.4	330	2 Q599J7_XENLA	Q599j7 xenopus lae
11	23	7.1	327	2 Q5CB07_BRARE	Q5cb07 brachydanio
12	22	6.8	303	2 Q599J9_ORYLA	Q599j9 oryzias lat
13	21	6.5	334	2 Q5CB02_BRARE	Q5cb02 brachydanio
14	21	6.5	335	2 Q5CB01_ORYLA	Q5cb01 oryzias lat
15	21	6.5	335	2 Q5CA29_TETNG	Q5ca29 tetraodon n
16	21	6.5	335	2 Q5CB00_FUGRU	Q5cb00 fugu rubrip
17	21	6.5	359	2 Q4SE14_TETNG	Q4se14 tetraodon n
18	20	6.2	304	2 Q5CA27_FUGRU	Q5ca27 fugu rubrip
19	20	6.2	304	2 Q5CA28_TETNG	Q5ca28 tetraodon n
20	20	6.2	316	2 Q6DBR2_BRARE	Q6db2 brachydanio
21	20	6.2	317	2 Q7T180_CHICK	Q7t180 gallus gall
22	20	6.2	413	2 Q4SG61_TETNG	Q4sg61 tetraodon n
23	19	5.9	333	2 Q599K2_BOVIN	Q599k2 bos taurus
24	15	4.6	277	2 Q8R1V2_MOUSE	Q8r1v2 mus musculu
25	15	4.6	317	2 Q7T1E1_CHICK	Q7t1e1 gallus gall
26	15	4.6	332	1 B3GA1_PANTR	Q5cb04 p galactosy
27	15	4.6	334	1 B3GA1_HUMAN	Q9p2w7 homo sapien
28	15	4.6	334	1 B3GA1_MOUSE	Q9cwf3 mus musculu
29	15	4.6	334	1 B3GA1_RAT	Q35789 r galactosy
30	15	4.6	334	2 Q96F87_HUMAN	Q96f87 homo sapien
31	15	4.6	334	2 Q6PIG8_MOUSE	Q6pig8 mus musculu

32	15	4.6	335	1 B3GA1_CANFA	Q5cb03 c galactosy
33	15	4.6	345	2 Q8BLT8_MOUSE	Q8blt8 mus musculu
34	15	4.6	347	2 Q8R531_MOUSE	Q8r531 mus musculu
35	15	4.6	394	2 Q5DTF9_MOUSE	Q5dtf9 mus musculu
36	14	4.3	334	2 Q5CB06_ORYLA	Q5cb06 oryzias lat
37	14	4.3	342	2 Q63ZM6_XENLA	Q63zm6 xenopus lae
38	13	4.0	159	2 Q8AYG7_CHICK	Q8ayg7 gallus gall
39	13	4.0	190	2 Q8ND18_HUMAN	Q8nd18 homo sapien
40	13	4.0	328	2 Q567G3_BRARE	Q567g3 brachydanio
41	13	4.0	329	2 Q8JH19_BRARE	Q8jhi9 brachydanio
42	13	4.0	339	2 Q599K3_XENTR	Q599k3 xenopus tro
43	12	3.7	303	2 Q599K4_XENTR	Q599k4 xenopus tro
44	11	3.4	148	2 Q8AYG6_CHICK	Q8ayg6 gallus gall
45	11	3.4	242	2 Q7T1E2_CHICK	Q7t1e2 gallus gall

#### ALIGNMENTS

RESULT 1  
B3GA2\_HUMAN  
ID B3GA2\_HUMAN STANDARD; PRT; 323 AA.  
AC QNPZ5; Q5JS09; Q8TF38; Q96NK4;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Galactosylgalactosylxylosylprotein 3-beta-glucuronosyltransferase 2  
DE (EC 2.4.1.135) (Beta-1,3-glucuronosyltransferase 2)  
DE (Glucuronosyltransferase-S) (GICAT-S) (UDP-glucuronosyltransferase-S)  
DE (GICAT-D).  
GN Names=B3GAT2; Synonyms=GLCATS, KIAA1963;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22410182; PubMed=12522689; DOI=10.1007/s100380200103;  
RA Marcos I., Galan J.J., Borrego S., Antinolo G.;  
RT "Cloning, characterization, and chromosome mapping of the human GICAT-S gene."  
RL J. Hum. Genet. 47:677-680 (2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
RC TISSUE=Brain;  
RX MEDLINE=21842142; PubMed=11853319;  
RA Nagase T., Kikuno R., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XXII. The complete sequences of 50 new cDNA clones which code for large proteins."  
RT DNA Res. 8:319-327 (2001).  
RN [3]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RX MEDLINE=22935783; PubMed=14574404; DOI=10.1038/nature02055;  
RA Mungall A.J., Palmer S.A., Sims S.K., Edwards C.A., Ashurst J.L., Wilming L., Jones M.C., Horton R., Hunt S.E., Scott C.E., Gilbert J.G.R., Clamp M.E., Bethel G., Milne S., Ainscough R., Almeida J.P., Ambrose K.D., Andrews T.D., Ashwell R.I.S., Babbage A.K., Bagguley C.L., Bailey J., Banerjee R., Barker D.J., Barlow K.F., Bates K., Beare D.M., Beasley H., Beasley O., Bird C.P., Blakey S.E., Bray-Allen S., Brook J., Brown A.J., Brown J.Y., Burford D.C., Burrill W., Burton J., Carder C., Carter N.P., Chapman J.C., Clark S.Y., Clark G., Clee C.M., Clegg S., Cobley V., Collier R.E., Collins J.E., Colman L.K., Corby N.R., Coville G.J., Culley K.M., Dhami P., Davies J., Dunn M., Earthrowl M.E., Ellington A.E., Evans K.A., Faulkner L., Francis M.D., Frankish A., Frankland J., French L., Garner P., Garnett J., Ghorri M.J., Gilby L.M., Gilson C.J., Glithero R.J., Graham D.V., Grant M., Grigble S., Griffiths C., Griffiths M.N.D., Hall R., Halls K.S., Hammond S., Harley J.L., Hart E.A., Heath P.D., Heathcote R., Holmes S.J., Howden P.J., Howe K.L., Howell G.R., Huckle E., Humphray S.J., Humphries M.D., Hunt A.R., Johnson C.M., Joy A.A.,

RA Kay M., Keenan S.J., Kimberley A.M., King A., Laird G.K., Langford C.,  
RA Lawlor S., Leongamornlert D.A., Leversha M., Lloyd C.R., Lloyd D.M.,  
RA Loveland J.E., Lovell J., Martin S., Mashreghi-Mohammadi M.,  
RA Maslen G.L., Matthews L., McCann O.T., McLaren S.J., McLeay K.,  
RA McMurray A., Moore M.J.F., Mullikin J.C., Niblett D., Nickerson T.,  
RA Novik K.L., Oliver K., Overton-Larty E.K., Parker A., Patel R.,  
RA Pearce A.V., Peck A.I., Phillimore B.J.C.T., Phillips S., Plumb R.W.,  
RA Porter K.M., Ramsey Y., Ranby S.A., Rice C.M., Ross M.T., Searle S.W.,  
RA Sehra H.K., Sheridan E., Skuce C.D., Smith S., Smith M., Spraggon L.,  
RA Squares S.L., Steward C.A., Sycamore N., Tamiyn-Hall G., Tester J.,  
RA Theaker A.J., Thomas D.W., Thorpe A., Tracey A., Tromans A., Tubby B.,  
RA Wall M., Wallis J.M., West A.P., White S.S., Whitehead S.L.,  
RA Whittaker H., Wild A., Willey D.J., Wilmer T.E., Wood J.M., Wray P.W.,  
RA Wyatt J.C., Young L., Younger R.M., Bentley D.R., Coulson A.,  
RA "The DNA sequence and analysis of human chromosome 6.,"  
RT Nature 425:805-811(2003).  
RL (4)  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 157-323.  
RC TISSUE=Fetal brain;  
RX PubMed=14702039; DOI=10.1038/ng1285;  
RA Oka T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,  
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,  
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,  
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,  
RA Nagahara K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,  
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,  
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,  
RA Omura Y., Abe K., Kamihara K., Katsuma T., Sato K., Tanikawa M.,  
RA Yamazaki M., Ninomiya K., Iehibashi T., Yamashita H., Murakawa K.,  
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,  
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hottuta T.,  
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,  
RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,  
RA Inose N., Musashino K., Yuki H., Oshima A., Sasaki N., Aotsuka S.,  
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,  
RA Moriyama S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,  
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,  
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,  
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,  
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,  
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,  
RA Kawabata A., Hakiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,  
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,  
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,  
RA Mizushima-Sugano J., Sato T., Shirai Y., Takahashi Y., Nakagawa K.,  
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,  
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;  
RT "Complete sequencing and characterization of 21,243 full-length human  
cDNAs.,"  
RL Nat. Genet. 36:40-45(2004).  
CC -I- FUNCTION: Involved in the biosynthesis of L2/HNK-1 carbohydrate  
CC epitope on both glycolipids and glycoproteins (By similarity).  
CC -I- CATALYTIC ACTIVITY: UDP-glucuronate + 3-beta-D-galactosyl-4-beta-  
CC D-galactosyl-O-beta-D-xylosylprotein = UDP + 3-beta-D-  
CC glucuronosyl-3-beta-D-galactosyl-4-beta-D-galactosyl-O-beta-D-  
CC xylosylprotein.  
CC -I- COFACTOR: Manganese.  
CC -I- PATHWAY: Glycosylation.  
CC -I- SUBUNIT: Homodimer (Potential).  
CC -I- SUBCELLULAR LOCATION: Type II membrane protein. Golgi.  
CC -I- TISSUE SPECIFICITY: Expressed in the trachea, retina, spinal cord,  
CC hippocampus and other brain regions, and, at lower levels, in  
CC testis and ovary.  
CC -I- SIMILARITY: Belongs to the glycosyltransferase 43 family.  
CC  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC

DR EMBL: AY070019; AAL57718.1; -; mRNA.  
DR EMBL: AY070110; AAL58977.1; -; Genomic DNA.  
DR EMBL: AY070108; AAL58977.1; JOINED; Genomic DNA.  
DR EMBL: AY070109; AAL58977.1; JOINED; Genomic DNA.  
DR EMBL: AB075843; BAB85549.1; -; ALT INIT; mRNA.  
DR EMBL: AL121961; CA142145.1; -; Genomic DNA.  
DR EMBL: AL450320; CA142145.1; JOINED; Genomic DNA.  
DR EMBL: AL450320; CA139582.1; -; Genomic DNA.  
DR EMBL: AL121961; CA139582.1; JOINED; Genomic DNA.  
DR EMBL: AK055248; BAB70889.1; ALT\_INIT; mRNA.  
DR HSSP: O94766; 1KNS.  
DR Ensemble: ENSG00000112309; Homo sapiens.  
DR HGNC: HGNC:922; B3GAT2.  
DR H-invDB: HIX0006000; -.  
DR MIM: 607497; -.  
DR GO: GO:0016021; C:integral to membrane; ISS.  
DR GO: GO:0015018; F:galactosylgalactosylxylosylprotein 3-beta-g. .; ISS.  
DR GO: GO:0016757; F:transferase activity, transferring glycosyl. .; ISS.  
DR GO: GO:0016051; P:carbohydrate biosynthesis; ISS.  
DR InterPro: IPR005027; Glyco\_trans\_43.  
DR PANTHER: PTHR10896; Glyco\_transf\_43; 1.  
DR Pfam: PF03360; Glyco\_transf\_43; 1.  
DR Glycoprotein; Golgi stack; Manganese; Metal-binding; Multigene family;  
KW Signal-anchor; Transferase; Transmembrane.  
FT TOPO\_DOM 1 2 Cytoplasmic (Potential).  
FT TRANSMEM 3 23 Signal-anchor for type II membrane  
FT TOPO\_DOM 24 323 Lumenal (Potential).  
FT ACT\_SITE 273 273 Proton acceptor (By similarity).  
FT METAL 187 187 Manganese (By similarity).  
FT CARBOHYD 67 67 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 292 292 N-linked (GlcNAc...) (Potential).  
SQ SEQUENCE 323 AA; 36919 MW; 85058D52D2D28463 CRC64;  
  
Query Match 100.0%; Score 323; DB 1; Length 323;  
Best Local Similarity 100.0%; Pred. No. 9.6e-312;  
Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MKSALFTRFFILLPWILVIMLDVDRPVPPLTPRPFPSYAVGRGARLP LRRGGPA 60  
DB 1 MKSALFTRFFILLPWILVIMLDVDRPVPPLTPRPFPSYAVGRGARLP LRRGGPA 60  
  
QY 61 HGTQKNSRQPOPEPQLPTIYATPTYSRVRQKELTRLANTPRVAQLHWILVEDAA 120  
DB 61 HGTQKNSRQPOPEPQLPTIYATPTYSRVRQKELTRLANTPRVAQLHWILVEDAA 120  
  
QY 121 ARSELVSRFLARAGLPSTHLHVPTRRYKRPGLPRATEORNAGLAWLRHORHQAQPGV 180  
DB 121 ARSELVSRFLARAGLPSTHLHVPTRRYKRPGLPRATEORNAGLAWLRHORHQAQPGV 180  
  
QY 181 LFFADDNDNTYSLELFQEMRTTRKVSVPVGLVGGRRYERPLVGVGVGYTGWRRDRPF 240  
DB 181 LFFADDNDNTYSLELFQEMRTTRKVSVPVGLVGGRRYERPLVGVGVGYTGWRRDRPF 240  
  
QY 241 AIDMAGFAVSLQVILSNPKAVFKRSGQMGQESFLKQITTVVEELPKANNCTKVLVWH 300  
DB 241 AIDMAGFAVSLQVILSNPKAVFKRSGQMGQESFLKQITTVVEELPKANNCTKVLVWH 300  
  
QY 301 TRTEKVNLANEPKYHLDVTVKIEV 323  
DB 301 TRTEKVNLANEPKYHLDVTVKIEV 323  
  
RESULT 2  
Q599K0 PANTR PRELIMINARY; PRT; 323 AA.  
AC Q599K0;  
DT 10-MAY-2005 (TrEMBLrel. 30, Created)  
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)  
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)  
DE Beta-3-glucuronyltransferase-S (EC 2.4.1.135).  
GN Name-b3gat2;  
OS Pan troglodytes (Chimpanzee).



QY 173 HORAQPCVLFPADDDNTYSLFQEMRTTRKVSVPVGLVGGRRYERPLVNGKVVGWYT 232  
Db 176 HORAQPCVLFPADDDNTYSLFQEMRTTRKVSVPVGLVGGRRYERPLVNGKVVGWYT 235

QY 233 GWRADRFALDMAGFAVSLQVILSNPKAVFKRGSGPQMGSDFL 277  
Db 236 GWRADRFALDMAGFAVSLQVILSNPKAVFKRGSGPQMGSDFL 280

## RESULT 5

QSCAS7\_CHICK  
ID QSCAS7\_CHICK PRELIMINARY; PRT; 304 AA.  
AC QSCAS7  
DT 10-MAY-2005 (TEMBLrel. 30, Created)  
DT 10-MAY-2005 (TEMBLrel. 30, Last sequence update)  
DT 10-MAY-2005 (TEMBLrel. 30, Last annotation update)  
DE Beta3-glucuronosyltransferase-S (EC 2.4.1.135).  
GN Name=B3GAT2;  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Ouzzine M., Magdalou J., Fournel-Gigleux S., Mollicone R., Oriol R.;  
RT "Phylogeny of beta3-glucuronosyltransferases.";  
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ889606; CAI:63864.1; -; mRNA.  
DR InterPro: IPR005027; Glyco\_trans\_43.  
DR PANTHER: PTHR10896; Glyco\_trans\_43; 6.  
DR Pfam: PF03360; Glyco\_transf\_43; 1.  
KW Glycosyltransferase; Transferase.  
SQ SEQUENCE 304 AA; 34618 MW; 5C45DF677FBEA624 CRC64;

Query Match 19.2%; Score 62; DB 2; Length 304;  
Best Local Similarity 100.0%; Pred. No. 1.6e-52;  
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 258 PKAVFKRRGSPQMGESDFLKQITVVEELPKANNCTKVLVHTRTEKVNLANEPKYHLD 317  
Db 239 PKAVFKRRGSPQMGESDFLKQITVVEELPKANNCTKVLVHTRTEKVNLANEPKYHLD 298  
QY 318 TV 319  
Db 299 TV 300

## RESULT 6

B3GA2\_MOUSE  
ID B3GA2\_MOUSE STANDARD; PRT; 324 AA.  
AC P59270;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Galactosylgalactosylxylosylprotein 3-beta-glucuronosyltransferase 2  
DE (EC 2.4.1.135) (Beta-1,3-glucuronosyltransferase 2)  
DE (Glucuronosyltransferase-S) (GLCAT-S) (UDP-glucuronosyltransferase-S)  
DE (GLCAT-D).  
GN Name=B3gat2; Synonyms=Glcats;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muroidae; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).  
RC STRAIN=BALB/c; TISSUE=Brain;  
RX MEDLINE=22271538; PubMed=12383500; DOI=10.1016/S0378-1119(02)00840-5;  
RA Imai K., Iehizaki T., Seiki T., Saito F., Inazawa J., Oka S.,  
RA Kawasaki T.;  
RT "cDNA cloning, genomic structure and chromosomal mapping of the mouse

RT glucuronosyltransferase-S involved in the biosynthesis of the HNK-1  
RT carbohydrate epitope.";  
RL Gene 296:29-36(2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 2).  
RC STRAIN=C57BL/6J; TISSUE=Embryonic head, and Kidney;  
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
RA Nikiishi I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomic M.,  
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
RA Yuan Z., Zvolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
RA Hara A., Hashizume W., Imotani K., Iehi Y., Itoh M., Kagawa I.,  
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
RA Birney E., Hayashizaki Y.;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
RN [3]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).  
RC STRAIN=C57BL/6; TISSUE=Brain;  
RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullen P.H.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -!- FUNCTION: Involved in the biosynthesis of L2/HNK-1 carbohydrate  
CC epitope on both glycolipids and glycoproteins.  
CC -!- CATALYTIC ACTIVITY: UDP-glucuronate + 3-beta-D-galactosyl-4-beta-  
CC D-galactosyl-O-beta-D-xylosylprotein + 3-beta-D-galactosyl-4-beta-D-  
CC glucuronosyl-3-beta-D-galactosyl-4-beta-D-galactosyl-O-beta-D-  
CC xylosylprotein.  
CC -!- COFACTOR: Manganese.  
CC -!- PATHWAY: Glycosylation.  
CC -!- SUBUNIT: Homodimer (potential).  
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (By  
CC similarity).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event-Alternative splicing; Named isoforms=2;

```
CC Name=1;
CC IsoId=P59270-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P59270-2; Sequence=VSP_001796, VSP_001797;
CC Note=No experimental confirmation available;
CC -1- TISSUE SPECIFICITY: Expressed in brain, but not in liver and
CC kidney.
CC -1- SIMILARITY: Belongs to the glycosyltransferase 43 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: AB055902; BAC20343.1; -; mRNA.
CC EMBL: AK048146; BAC33257.1; -; mRNA.
CC EMBL: AK052640; BAC35075.1; -; mRNA.
CC EMBL: BC056368; AAH56368.1; -; mRNA.
CC EMBL: BC058082; AAH58082.1; -; mRNA.
CC HSP: O94766; 1KWS.
CC Ensembl: ENSMUSG0000026156; Mus musculus.
CC MGI: MGI:2389490; B3gat2.
CC GO: GO:0016021; C: integral to membrane; ISS.
CC GO: GO:0015018; F: galactosylgalactosylxylosylprotein 3-beta-g. .; ISS.
CC GO: GO:0016757; F: transferase activity, transferring glycosyl. .; ISS.
CC GO: GO:0016051; P: carbohydrate biosynthesis; ISS.
CC InterPro: IPR005027; Glyco trans 43.
CC PANTHER: PTHR10896; Glyco trans 43; 1.
CC Pfam: PF03360; Glyco trans 43; 1.
CC Alternative splicing: Glycoprotein; Golgi stack; Manganese;
KW Metal-binding; Multigene family; Signal-anchor; Transferase;
KW Transmembrane.
FT TOPO_DOM 1 2 Cytoplasmic (Potential).
FT TRANSMEM 3 23 Signal-anchor for type II membrane
FT FT
FT TOPO_DOM 24 324 Lumenal (Potential).
FT ACT_SITE 274 324 Proton acceptor (By similarity).
FT METAL 188 188 Manganese (By similarity).
FT CARBOHYD 68 68 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 293 293 N-linked (GlcNAc...) (Potential).
FT VARSPPLIC 248 302 FAVSLQVILSNPKAVFKRRGSGPQMGSDFLKQITVVELEP
PKASNCTKVLVWHT -> EQNADVICPRGMGRNWERPIF
IKPLSVLHSSEILLPKIGSLWQVEVDFRSG (in
isoform 2).
FT FTId=VSP_001796.
FT Missing (in isoform 2).
FT VARSPPLIC 303 324 /FTId=VSP_001797.
FT CONFLICT 291 291 A -> V (in Ref. 1).
FT SEQUENCE 324 AA; 37132 MW; 675C549D3787B957 CRC64;

Query Match 16.7%; Score 54; DB 1; Length 324;
Best Local Similarity 100.0%; Pred. No. 1.5e-44;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 237 DRPFAIDMAGFAVSLQVILSNPKAVFKRRGSGPQMGSDFLKQITVVELEPKA 290
DB 238 DRPFAIDMAGFAVSLQVILSNPKAVFKRRGSGPQMGSDFLKQITVVELEPKA 291
|||||
RESULT 7
QSDTT7 MOUSE PRELIMINARY; PRT; 349 AA.
AC QSDTT7;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE MKTAA1963 protein (Fragment)
GN Name=B3gat2; Synonyms=MKIA1963;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
```

```
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Retal brain;
RA Okazaki N., Kikuno R.F., Ohara R., Inamoto S., Nagase T., Ohara O.,
RA Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of KIAA Gene.
RT The Complete Nucleotide Sequences of Mouse KIAA-homologous cDNAs
RT Identified by Screening of Terminal sequences of cDNA Clones Randomly
RT Sampled from Size-Fractionated Libraries. ";
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK220433; BAD90265.1; -; mRNA.
DR MGI: MGI:2389490; B3gat2.
DR GO: GO:0015020; F: glucuronosyltransferase activity; TAS.
DR InterPro: IPR005027; Glyco trans 43.
DR Pfam: PF03360; Glyco trans 43; 1.
FT NON_TER 1
SQ SEQUENCE 349 AA; 40066 MW; 2EFB2B8BCBD84169 CRC64;

Query Match 16.7%; Score 54; DB 2; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.6e-44;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 237 DRPFAIDMAGFAVSLQVILSNPKAVFKRRGSGPQMGSDFLKQITVVELEPKA 290
DB 263 DRPFAIDMAGFAVSLQVILSNPKAVFKRRGSGPQMGSDFLKQITVVELEPKA 316
|||||
RESULT 8
B3GA2 RAT STANDARD; PRT; 324 AA.
ID B3GA2 RAT
AC Q92137;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Galactosylgalactosylxylosylprotein 3-beta-glucuronosyltransferase 2
DE (EC 2.4.1.135) (Beta-1,3-glucuronosyltransferase 2)
DE (Glucuronosyltransferase-S) (GlcAT-S) (UDP-glucuronosyltransferase-S)
DE (GlcAT-D).
GN Name=B3gat2; Synonyms=Glcats;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=99185317; PubMed=10082676; DOI=10.1006/bbrc.1999.0151;
RA Seiki T., Oka S., Terayama K., Imiya K., Kawasaki T.;
RT "Molecular cloning and expression of a second glucuronosyltransferase
RT involved in the biosynthesis of the HNK-1 carbohydrate epitope. ";
RL Biochem. Biophys. Res. Commun. 255:182-187(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE, AND CHARACTERIZATION.
RC TISSUE=Embryonic brain;
RX MEDLINE=99287914; PubMed=10358066; DOI=10.1074/jbc.274.24.17115;
RA Shimoda Y., Tajima Y., Nagase T., Harii K., Osumi N., Sanai Y.;
RT "Cloning and expression of a novel galactoside beta1,3-
RT glucuronyltransferase involved in the biosynthesis of HNK-1 epitope. ";
RL J. Biol. Chem. 274:17115-17122(1999).
CC -1- FUNCTION: Involved in the biosynthesis of L2/HNK-1 carbohydrate
CC epitope on both glycolipids and glycoproteins. Substrates include
CC asialo-orosomucoid (ASOR), paragloboside (lacto-N-
CC neotetraosylceramide), Gal-beta-1,4-GlcNAc-beta-1,3-Gal-beta-1,4-
CC Glc-pyridylamine and Gal-beta-1,3-GlcNAc-beta-1,3-Gal-beta-1,4-
CC Glc-pyridylamine.
CC -1- CATALYTIC ACTIVITY: UDP-glucuronate + 3-beta-D-galactosyl-4-beta-
CC D-galactosyl-O-beta-D-xylosylprotein = UDP + 3-beta-D-
CC glucuronosyl-3-beta-D-galactosyl-4-beta-D-galactosyl-O-beta-D-
CC xylosylprotein.
CC -1- COFACTOR: Manganese.
CC -1- PATHWAY: Glycosylation.
```



```
RESULT 12
Q599J9_ORYLA
ID Q599J9_ORYLA PRELIMINARY; PRT; 303 AA.
AC Q599J9;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Beta-3-glucuronyltransferase-S (EC 2.4.1.135).
GN Name=b3gat2;
OS Oryzias latipes (Medaka fish) (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=8090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ouzzine M., Magdalou J., Fournel-Gigleux S., Mollicone R., Oriol R.;
RT "Phylogeny of beta3-glucuronyltransferases.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ890842; CA168028.1; -; mRNA.
DR InterPro; IPR011990; TPR-like_helical.
DR PANTHER; PTHR10896; Glyco_trans_43; 7.
DR Pfam; PF03360; Glyco_transf_43; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 303 AA; 33932 MW; 6CD039E7361589A7 CRC64;

Query Match 6.8%; Score 22; DB 2; Length 303;
Best Local Similarity 100.0%; Pred. No. 8.4e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 AITPTYSRPVQKAELTRLNTF 105
|
Db 67 AITPTYSRPVQKAELTRLNTF 88

RESULT 13
Q5CB02_BRARE
ID Q5CB02_BRARE PRELIMINARY; PRT; 334 AA.
AC Q5CB02;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Beta3-glucuronyltransferase (EC 2.4.1.135).
GN Name=b3gat1;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ouzzine M., Magdalou J., Fournel-Gigleux S., Mollicone R., Oriol R.;
RT "Phylogeny of beta3-glucuronyltransferases.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ898976; CA162041.1; -; mRNA.
DR InterPro; IPR005027; Glyco_trans_43.
DR InterPro; IPR011990; TPR-like_helical.
DR PANTHER; PTHR10896; Glyco_trans_43; 7.
DR Pfam; PF03360; Glyco_transf_43; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 334 AA; 38370 MW; 4F5556173941A4C2 CRC64;

Query Match 6.5%; Score 21; DB 2; Length 334;
Best Local Similarity 100.0%; Pred. No. 8.9e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 85 ITPTYSRPVQKAELTRLNTF 105
|
Db 89 ITPTYSRPVQKAELTRLNTF 109

RESULT 14
Q5CB01_ORYLA
ID Q5CB01_ORYLA PRELIMINARY; PRT; 335 AA.
```

```
Q5CB01;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Beta3-glucuronyltransferase (EC 2.4.1.135).
GN Name=b3gat1;
OS Oryzias latipes (Medaka fish) (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=8090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ouzzine M., Magdalou J., Fournel-Gigleux S., Mollicone R., Oriol R.;
RT "Phylogeny of beta3-glucuronyltransferases.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ888977; CA162042.1; -; mRNA.
DR InterPro; IPR005027; Glyco_trans_43.
DR InterPro; IPR011990; TPR-like_helical.
DR PANTHER; PTHR10896; Glyco_trans_43; 7.
DR Pfam; PF03360; Glyco_transf_43; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 335 AA; 38244 MW; 365B1E0E76C10B31 CRC64;

Query Match 6.5%; Score 21; DB 2; Length 335;
Best Local Similarity 100.0%; Pred. No. 9e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 85 ITPTYSRPVQKAELTRLNTF 105
|
Db 90 ITPTYSRPVQKAELTRLNTF 110

RESULT 15
Q5CAZ9_TETNG
ID Q5CAZ9_TETNG PRELIMINARY; PRT; 335 AA.
AC Q5CAZ9;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Beta3-glucuronyltransferase (EC 2.4.1.135).
GN Name=b3gat1;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ouzzine M., Magdalou J., Fournel-Gigleux S., Mollicone R., Oriol R.;
RT "Phylogeny of beta3-glucuronyltransferases.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ888979; CA162044.1; -; mRNA.
DR InterPro; IPR005027; Glyco_trans_43.
DR InterPro; IPR011990; TPR-like_helical.
DR PANTHER; PTHR10896; Glyco_trans_43; 7.
DR Pfam; PF03360; Glyco_transf_43; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 335 AA; 38292 MW; F89E348F901C5A25 CRC64;

Query Match 6.5%; Score 21; DB 2; Length 335;
Best Local Similarity 100.0%; Pred. No. 9e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 85 ITPTYSRPVQKAELTRLNTF 105
|
Db 90 ITPTYSRPVQKAELTRLNTF 110

Search completed: May 9, 2006, 10:58:11
```

Job time : 60 secs

---

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2006, 10:57:17 ; Search time 27 Seconds  
(without alignments)  
989.046 Million cell updates/sec

Title: US-10-634-905-2  
Perfect score: 323  
Sequence: 1 MKSALFTRFFILLPWILVI.....EKNVLANEPKYHLDTVKIEV 323

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 572060 seqs, 82675679 residues

Word size : 1

Total number of hits satisfying chosen parameters: 570988

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/6\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/H\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/PCUTUS\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/RE\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	323	100.0	323	2	US-09-816-095-2
2	197	61.0	197	2	US-09-816-095-5
3	47	14.6	324	2	US-09-816-095-4
4	15	4.6	347	2	US-09-059-369-2
5	11	3.4	22	2	US-09-059-369-5
6	10	3.1	63	2	US-09-621-976-7125
7	8	2.5	25	2	US-09-059-369-4
8	2.5	107	2	2	US-09-328-352-5562
9	8	2.5	142	2	US-09-270-767-45263
10	8	2.5	256	2	US-09-252-991A-25670
11	8	2.5	405	2	US-09-252-991A-16887
12	8	2.5	410	2	US-09-252-991A-30570
13	8	2.5	530	2	US-09-252-991A-23339
14	8	2.5	2628	2	US-09-413-814-11
15	7	2.2	80	2	US-09-621-976-6610
16	7	2.2	90	2	US-09-949-002-438
17	7	2.2	117	2	US-09-270-767-41066
18	7	2.2	117	2	US-09-270-767-56282
19	7	2.2	131	2	US-09-902-540-11259
20	7	2.2	135	2	US-09-489-039A-10174
21	7	2.2	172	2	US-09-252-991A-25750
22	7	2.2	181	2	US-09-605-703B-1344
23	7	2.2	208	2	US-09-252-991A-27197
24	7	2.2	237	1	US-08-103-998-4
25	7	2.2	249	2	US-09-328-352-5813
26	7	2.2	254	2	US-09-902-540-16721
27	7	2.2	257	2	US-09-252-991A-21740

28	7	2.2	263	2	US-10-104-047-2192	Sequence 2192, Ap
29	7	2.2	230	2	US-09-543-681A-7078	Sequence 7078, Ap
30	7	2.2	291	2	US-09-270-767-33192	Sequence 33192, A
31	7	2.2	291	2	US-09-270-767-48409	Sequence 48409, A
32	7	2.2	304	2	US-09-403-343B-22	Sequence 22, Appl
33	7	2.2	310	2	US-09-252-991A-27488	Sequence 27488, A
34	7	2.2	310	2	US-09-328-352-7912	Sequence 7912, Ap
35	7	2.2	328	2	US-09-252-991A-26345	Sequence 26345, A
36	7	2.2	333	2	US-09-252-991A-19956	Sequence 19956, A
37	7	2.2	343	1	US-08-180-209B-56	Sequence 56, Appl
38	7	2.2	343	2	US-08-474-853-56	Sequence 56, Appl
39	7	2.2	343	2	US-09-166-205B-56	Sequence 56, Appl
40	7	2.2	343	4	PCT-US94-02629-56	Sequence 56, Appl
41	7	2.2	347	2	US-09-806-658-10	Sequence 10, Appl
42	7	2.2	357	2	US-09-252-991A-22404	Sequence 22404, A
43	7	2.2	358	2	US-09-710-279-18	Sequence 18, Appl
44	7	2.2	362	2	US-09-134-001C-5403	Sequence 5403, Ap
45	7	2.2	390	2	US-09-902-540-12817	Sequence 12817, A

ALIGNMENTS

RESULT 1  
US-09-816-095-2  
; Sequence 2, Application US/09816095  
; Patent No. 6664084  
; GENERAL INFORMATION:  
; APPLICANT: GAN, Weiniu  
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES  
; FILE REFERENCES: CL001147  
; CURRENT APPLICATION NUMBER: US/09/816,095  
; CURRENT FILING DATE: 2001-03-26  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 323  
; TYPE: PRT  
; ORGANISM: Human  
; US-09-816-095-2

Query Match	100.00%;	Score 323;	DB 2;	Length 323;
Best Local Similarity	100.00%;	Pred. No. 1.4e-316;		
Matches 323;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MKSALFTRFFILLPWILVIIMLDVDTTRRPVPLTPRPYSPYAVGGGARLPLRRGGPA	60	
Db	1	MKSALFTRFFILLPWILVIIMLDVDTTRRPVPLTPRPYSPYAVGGGARLPLRRGGPA	60	
Qy	61	HGTQKNQSRPQPPFPPLTIYAITPTYSRPVKAEITRLANTFRQVAQLHWILVEDAA	120	
Db	61	HGTQKNQSRPQPPFPPLTIYAITPTYSRPVKAEITRLANTFRQVAQLHWILVEDAA	120	
Qy	121	ARSELVSRFLARAGLSTHLHVTTPRYKRPGLPRATEQORNAGLWLRQHRQHQAOPGV	180	
Db	121	ARSELVSRFLARAGLSTHLHVTTPRYKRPGLPRATEQORNAGLWLRQHRQHQAOPGV	180	
Qy	181	LFADDNTYSLFLFQEMRTTRKVSVPVGLVGGRYERPLVGVGVGTYTGWDRDRPF	240	
Db	181	LFADDNTYSLFLFQEMRTTRKVSVPVGLVGGRYERPLVGVGVGTYTGWDRDRPF	240	
Qy	241	AIDMAGFVSLQVILSNPKAVFKRGSGQSQMSDFLKQITTVVEELPKANNCTKVLVWH	300	
Db	241	AIDMAGFVSLQVILSNPKAVFKRGSGQSQMSDFLKQITTVVEELPKANNCTKVLVWH	300	
Qy	301	TRTEKVNLANEPKYHLDTVKIEV	323	
Db	301	TRTEKVNLANEPKYHLDTVKIEV	323	

RESULT 2

```
US-09-816-095-5
; Sequence 5, Application US/09816095
; Patent No. 6664084
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001147
; CURRENT APPLICATION NUMBER: US/09/816,095
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Human
US-09-816-095-5
Query Match 61.0%; Score 197; DB 2; Length 197;
Best Local Similarity 100.0%; Pred. No. 5.3e-190;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKSALTRFILLPWILVIMLDVTRRPVPLTPRPYFSPYAVGCGARLPRLRGGA 60
DB 1 MKSALTRFILLPWILVIMLDVTRRPVPLTPRPYFSPYAVGCGARLPRLRGGA 60
QY 61 HGTKRNQSRQPQPEPQLPTIYAITPTYSRPVQKAEIIRLANTFPRVAQLHWILVEDAA 120
DB 61 HGTKRNQSRQPQPEPQLPTIYAITPTYSRPVQKAEIIRLANTFPRVAQLHWILVEDAA 120
QY 121 ARSELVSRFLARAGLSTHLHVTPTRYPKPGIIPRATEGNAGLAWLRQHQRHQAQPGV 180
DB 121 ARSELVSRFLARAGLSTHLHVTPTRYPKPGIIPRATEGNAGLAWLRQHQRHQAQPGV 180
QY 181 LFFADDDNTYSLELFOE 197
DB 181 LFFADDDNTYSLELFOE 197
RESULT 3
US-09-816-095-4
; Sequence 4, Application US/09816095
; Patent No. 6664084
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001147
; CURRENT APPLICATION NUMBER: US/09/816,095
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-816-095-4
Query Match 14.6%; Score 47; DB 2; Length 324;
Best Local Similarity 100.0%; Pred. No. 7e-39;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 176 AQFGVLFFADDDNTYSLELFOEQMRTRTKVSWPVLGVGGRRYERPLV 222
DB 177 AQFGVLFFADDDNTYSLELFOEQMRTRTKVSWPVLGVGGRRYERPLV 223
RESULT 4
US-09-059-369-2
; Sequence 2, Application US/09059369
; Patent No. 6040156
; GENERAL INFORMATION:
; APPLICANT: KAWASAKI, TOSHISUKE
; TITLE OF INVENTION: DNA ENCODING GLUCURONYLTRANSFERASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/059,369
; FILING DATE: 14-APR-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 9-127065
; FILING DATE: 16-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 9378-0002-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-059-369-2
Query Match 4.6%; Score 15; DB 2; Length 347;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 86 TPTYSRPVQKAELTR 100
DB 103 TPTYSRPVQKAELTR 117
RESULT 5
US-09-059-369-5
; Sequence 5, Application US/09059369
; Patent No. 6040156
; GENERAL INFORMATION:
; APPLICANT: KAWASAKI, TOSHISUKE
; TITLE OF INVENTION: DNA ENCODING GLUCURONYLTRANSFERASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/059,369
```

; FILING DATE: 14-APR-1998  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 9-127065  
; FILING DATE: 16-MAY-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 9378-0002-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 22 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-059-369-5

Query Match 3.4%; Score 11; DB 2; Length 22;  
Best Local Similarity 100.0%; Pred. No. 0.00097;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 239 PFAIDMAGFAV 249  
Db 8 PFAIDMAGFAV 18  
|||||

## RESULT 6

US-09-621-976-7125  
; Sequence 7125, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 7125  
; LENGTH: 63  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-621-976-7125

Query Match 3.1%; Score 10; DB 2; Length 63;  
Best Local Similarity 100.0%; Pred. No. 0.026;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 296 VLVVHTRTEK 305  
Db 32 VLVVHTRTEK 41  
|||||

## RESULT 7

US-09-059-369-4  
; Sequence 4, Application US/09059369  
; Patent No. 6040156  
; GENERAL INFORMATION:  
; APPLICANT: KAWASAKI, TOSHISUKE  
; APPLICANT: OKA, SHOGO  
; TITLE OF INVENTION: DNA ENCODING GLUCURONYLTRANSFERASE  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCELLELAND, MAIER & NEUSTADT,  
; ADDRESS: P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR  
; CITY: ARLINGTON

; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/059,369  
; FILING DATE: 14-APR-1998  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 9-127065  
; FILING DATE: 16-MAY-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 9378-0002-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-059-369-4

Query Match 2.5%; Score 8; DB 2; Length 25;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 183 FADDDNTY 190  
Db 9 FADDDNTY 16  
|||||

## RESULT 8

US-09-328-352-5562  
; Sequence 5562, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 5562  
; LENGTH: 107  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-5562

Query Match 2.5%; Score 8; DB 2; Length 107;  
Best Local Similarity 100.0%; Pred. No. 4.4;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 96 AELTRLAN 103  
Db 71 AELTRLAN 78  
|||||

## RESULT 9

US-09-270-767-45263  
; Sequence 45263, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 45263  
; LENGTH: 142  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-45263

Query Match 2.5%; Score 8; DB 2; Length 142;  
Best Local Similarity 100.0%; Pred. No. 5.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 243 DMAGFAVS 250  
Db 54 DMAGFAVS 61  
|||||

## RESULT 10

US-09-252-991A-25670  
; Sequence 25670, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 25670  
; LENGTH: 256  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-25670

Query Match 2.5%; Score 8; DB 2; Length 256;  
Best Local Similarity 100.0%; Pred. No. 9.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 GGGARLP 53  
Db 24 GGGARLP 31  
|||||

## RESULT 11

US-09-252-991A-16887  
; Sequence 16887, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 16887  
; LENGTH: 405  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-16887

Query Match 2.5%; Score 8; DB 2; Length 405;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 HORAPGV 180  
Db 270 HORAPGV 277  
|||||

## RESULT 12

US-09-252-991A-30570  
; Sequence 30570, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 30570  
; LENGTH: 410  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-30570

Query Match 2.5%; Score 8; DB 2; Length 410;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 QRHQHRA 176  
Db 394 QRHQHRA 401  
|||||

## RESULT 13

US-09-252-991A-23339  
; Sequence 23339, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 23339  
; LENGTH: 530  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-23339

Query Match 2.5%; Score 8; DB 2; Length 530;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 VGLVGRR 216  
Db 463 VGLVGRR 470  
|||||

## RESULT 14

US-09-413-814-11  
; Sequence 11, Application US/09413814  
; Patent No. 6225064  
; GENERAL INFORMATION:  
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH  
; APPLICANT: Bristol-Myers Squibb, Co.  
; APPLICANT: Beyer, Stefan  
; APPLICANT: Bloecker, Helmut  
; APPLICANT: Brandt, Petra  
; APPLICANT: Cino, Paul M  
; APPLICANT: Dougherty, Brian A  
; APPLICANT: Goldberg, Steven L  
; APPLICANT: Hofle, Gerhard  
; APPLICANT: Mueller, Joachim  
; APPLICANT: Reichenbach, Hans  
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or  
; TITLE OF INVENTION: heteropolyketide compounds  
; FILE REFERENCE: PCT/US 99/23535  
; CURRENT APPLICATION NUMBER: US/09/413,814  
; CURRENT FILING DATE: 1999-10-07  
; EARLIER APPLICATION NUMBER: DE 198 46 493.2  
; EARLIER FILING DATE: 1998-10-09  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: Patentin ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 2628  
; TYPE: PRT  
; ORGANISM: Sorangium cellulosum  
US-09-413-814-11

Query Match 2.5%; Score 8; DB 2; Length 2628;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 VGRGGARL 52  
|||  
Db 2456 VGRGGARL 2463

## RESULT 15

US-09-621-976-6610  
; Sequence 6610, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 6610  
; LENGTH: 80  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-621-976-6610

Query Match 2.2%; Score 7; DB 2; Length 80;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 274 SDFLKQI 280  
|||  
Db 9 SDFLKQI 15

Search completed: May 9, 2006, 10:58:44  
Job time : 27 secs

**This Page Blank (uspto)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2006, 10:58:21 ; Search time 78 Seconds  
(without alignments)  
1730.242 Million cell updates/sec

Title: US-10-634-905-2  
Perfect score: 323  
Sequence: 1 MKSALTRFFILLPWILVI.....EKVNLANEKPYHLDTVKIEV 323

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1867569 seqs, 417829326 residues

Word size : 1  
Total number of hits satisfying chosen parameters: 1866650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA\_Main:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	323	100.0	323	3	US-09-816-095-2
2	323	100.0	323	4	US-10-634-905-2
3	197	61.0	197	3	US-09-816-095-5
4	197	61.0	197	4	US-10-634-905-5
5	81	25.1	145	4	US-10-080-254-93
6	81	25.1	145	4	US-10-242-355-637
7	47	14.6	324	3	US-09-816-095-4
8	47	14.6	324	4	US-10-634-905-4
9	13	4.0	40	4	US-10-029-386-33145
10	13	3.7	167	5	US-10-450-763-43912
11	11	3.4	335	4	US-10-698-190-16
12	11	3.4	335	5	US-10-723-860-4238
13	11	3.4	691	5	US-10-450-763-30691
14	10	3.1	201	4	US-10-403-571-34
15	8	2.5	57	4	US-10-425-115-346410
16	8	2.5	80	4	US-10-437-963-139231
17	8	2.5	87	4	US-10-425-115-313878
18	8	2.5	263	5	US-10-849-948-4
19	8	2.5	277	4	US-10-437-963-143548
20	8	2.5	276	4	US-10-437-963-119079
21	8	2.5	392	4	US-10-156-761-9175
22	8	2.5	402	5	US-10-732-923-7823
23	8	2.5	479	6	US-11-097-143-13308
24	8	2.5	612	4	US-10-437-963-139689
25	8	2.5	807	4	US-10-138-221-7
26	8	2.5	840	4	US-10-425-115-204488
27	8	2.5	912	5	US-10-631-467-1543

28	7	2.2	19	3	US-09-864-761-39306	Sequence 39306, A
29	7	2.2	36	3	US-09-864-761-41351	Sequence 41351, A
30	7	2.2	44	4	US-10-242-355-404	Sequence 404, App
31	7	2.2	57	3	US-09-729-674-38	Sequence 38, Appl
32	7	2.2	57	5	US-10-913-553-38	Sequence 38, Appl
33	7	2.2	58	4	US-10-425-115-295950	Sequence 295950, A
34	7	2.2	64	4	US-10-767-701-50698	Sequence 50698, A
35	7	2.2	70	4	US-10-425-115-275941	Sequence 275941, A
36	7	2.2	70	4	US-10-425-115-336480	Sequence 336480, A
37	7	2.2	72	4	US-10-424-599-183280	Sequence 183280, A
38	7	2.2	72	4	US-10-767-701-50207	Sequence 50207, A
39	7	2.2	76	4	US-10-424-599-183282	Sequence 183282, A
40	7	2.2	80	5	US-10-287-436A-603	Sequence 603, App
41	7	2.2	82	5	US-10-450-763-57491	Sequence 57491, A
42	7	2.2	86	4	US-10-425-115-200651	Sequence 200651, A
43	7	2.2	87	3	US-09-882-227-422	Sequence 422, App
44	7	2.2	87	4	US-10-335-977-7681	Sequence 7681, Ap
45	7	2.2	89	4	US-10-437-963-112770	Sequence 112770, A

ALIGNMENTS

RESULT 1  
US-09-816-095-2  
; Sequence 2, Application US/09816095  
; Patent No. US20020137164A1  
; GENERAL INFORMATION:  
; APPLICANT: GAN, Weiniu  
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001147  
; CURRENT APPLICATION NUMBER: US/09/816,095  
; CURRENT FILING DATE: 2001-03-26  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 323  
; TYPE: PRT  
; ORGANISM: Human  
US-09-816-095-2

Query Match		100.0%;	Score 323;	DB 3;	Length 323;
Best Local Similarity		100.0%;	Pred. No. 3.2e-300;		
Matches 323;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MKSALTRFFILLPWILVIIMLDVDRTRPPPLTPRPYFSPYAVGRCGARLPLRRGGPA	60		
DB	1	MKSALTRFFILLPWILVIIMLDVDRTRPPPLTPRPYFSPYAVGRCGARLPLRRGGPA	60		
QY	61	HGTQKRNQSRQPOPEPQLPTTIAITPTYSRPVQKAELTRLANTFRQVQLHILVEDAA	120		
DB	61	HGTQKRNQSRQPOPEPQLPTTIAITPTYSRPVQKAELTRLANTFRQVQLHILVEDAA	120		
QY	121	ARSELVSRFLARAGLPSTHLVPTPRRYKRGPLPRATEORNAGLAWLRQRHQRAQPGV	180		
DB	121	ARSELVSRFLARAGLPSTHLVPTPRRYKRGPLPRATEORNAGLAWLRQRHQRAQPGV	180		
QY	181	LFADFDDNTYSLELFOEMMTTRKVSVPVGLVGGRRYERPLVGVGVGMYTCWRADRP	240		
DB	181	LFADFDDNTYSLELFOEMMTTRKVSVPVGLVGGRRYERPLVGVGVGMYTCWRADRP	240		
QY	241	AIDMAGFVSLQVILSNPKAVFKRGSGQGMQESDFLKQITTVVEELPKANNCTKVLVWH	300		
DB	241	AIDMAGFVSLQVILSNPKAVFKRGSGQGMQESDFLKQITTVVEELPKANNCTKVLVWH	300		
QY	301	TRTEKVNLANEPKYHLDTVKIEV	323		
DB	301	TRTEKVNLANEPKYHLDTVKIEV	323		
RESULT 2					

```
US-10-634-905-2
; Sequence 2, Application US/10634905
; Publication No. US20040067225A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001147-DIV
; CURRENT FILING DATE: 2003-08-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-634-905-2

Query Match      100.0%; Score 323; DB 4; Length 323;
Best Local Similarity 100.0%; Pred. No. 3.2e-300;
Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 M K S A L T R F F I L L P W I L I V I M L D V T R R P V P P L T P R P Y F S P Y A V G R G G A R L P L R R G G P A 60
D b 1 M K S A L T R F F I L L P W I L I V I M L D V T R R P V P P L T P R P Y F S P Y A V G R G G A R L P L R R G G P A 60
QY 61 H G T Q K N Q S R P Q P Q P E P Q L P T I A I T P T Y S R P V Q K A E L T R L A N T F R Q V A Q L H W I L V E D A A 120
D b 61 H G T Q K N Q S R P Q P Q P E P Q L P T I A I T P T Y S R P V Q K A E L T R L A N T F R Q V A Q L H W I L V E D A A 120
QY 121 A R S E L V S R F L A R A G L P S T H L H V P T P R R Y K R P G L P R A T E O R N A G L A W L R Q R H O H Q R A O P G V 180
D b 121 A R S E L V S R F L A R A G L P S T H L H V P T P R R Y K R P G L P R A T E O R N A G L A W L R Q R H O H Q R A O P G V 180
QY 181 L F F A D D D N T Y S L E L F Q E 197
D b 181 L F F A D D D N T Y S L E L F Q E 197

RESULT 4
US-10-634-905-5
; Sequence 5, Application US/10634905
; Publication No. US20040067225A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001147-DIV
; CURRENT APPLICATION NUMBER: US/10/634,905
; CURRENT FILING DATE: 2003-08-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-634-905-5

Query Match      61.0%; Score 197; DB 4; Length 197;
Best Local Similarity 100.0%; Pred. No. 7.3e-180;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 M K S A L T R F F I L L P W I L I V I M L D V T R R P V P P L T P R P Y F S P Y A V G R G G A R L P L R R G G P A 60
D b 1 M K S A L T R F F I L L P W I L I V I M L D V T R R P V P P L T P R P Y F S P Y A V G R G G A R L P L R R G G P A 60
QY 61 H G T Q K N Q S R P Q P Q P E P Q L P T I A I T P T Y S R P V Q K A E L T R L A N T F R Q V A Q L H W I L V E D A A 120
D b 61 H G T Q K N Q S R P Q P Q P E P Q L P T I A I T P T Y S R P V Q K A E L T R L A N T F R Q V A Q L H W I L V E D A A 120
QY 121 A R S E L V S R F L A R A G L P S T H L H V P T P R R Y K R P G L P R A T E O R N A G L A W L R Q R H O H Q R A O P G V 180
D b 121 A R S E L V S R F L A R A G L P S T H L H V P T P R R Y K R P G L P R A T E O R N A G L A W L R Q R H O H Q R A O P G V 180
QY 181 L F F A D D D N T Y S L E L F Q E 197
D b 181 L F F A D D D N T Y S L E L F Q E 197

RESULT 5
US-10-080-254-93
; Sequence 93, Application US/10080254
; Publication No. US20030199008A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAIL14C1
; CURRENT APPLICATION NUMBER: US/10/080,254
; CURRENT FILING DATE: 2002-02-22
; NUMBER OF SEQ ID NOS: 169
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 93
; LENGTH: 145
; TYPE: PRT

QY 1 M K S A L T R F F I L L P W I L I V I M L D V T R R P V P P L T P R P Y F S P Y A V G R G G A R L P L R R G G P A 60
```

```
; ORGANISM: Homo sapiens
US-10-080-254-93

Query Match      25.1%; Score 81; DB 4; Length 145;
Best Local Similarity 100.0%; Pred. No. 5.4e-69;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 210 GLVGGRRYERPLVNGKVGVGTYGTWRADRPFAIDMAGFAVSLQVILSNPKAVFKRRGSQP 269
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 23  GLVGGRRYERPLVNGKVGVGTYGTWRADRPFAIDMAGFAVSLQVILSNPKAVFKRRGSQP 82

Qy 270 GMOESDFLKQITTVEELEPKA 290
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 83  GMOESDFLKQITTVEELEPKA 103

RESULT 6
US-10-242-355-637
; Sequence 637, Application US/10242355
; Publication No. US20030235831A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC003C1
; CURRENT APPLICATION NUMBER: US/10/242,355
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,897
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 637
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-242-355-637

Query Match      25.1%; Score 81; DB 4; Length 145;
Best Local Similarity 100.0%; Pred. No. 5.4e-69;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 210 GLVGGRRYERPLVNGKVGVGTYGTWRADRPFAIDMAGFAVSLQVILSNPKAVFKRRGSQP 269
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 23  GLVGGRRYERPLVNGKVGVGTYGTWRADRPFAIDMAGFAVSLQVILSNPKAVFKRRGSQP 82

Qy 270 GMOESDFLKQITTVEELEPKA 290
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 83  GMOESDFLKQITTVEELEPKA 103

RESULT 7
US-09-816-095-4
; Sequence 4, Application US/09816095
; Patent No. US20020137164A1
; GENERAL INFORMATION:
```

```
; APPLICANT: GAN, Weiniu
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; FILE REFERENCE: CL001147
; CURRENT APPLICATION NUMBER: US/09/816,095
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-816-095-4

Query Match      14.6%; Score 47; DB 3; Length 324;
Best Local Similarity 100.0%; Pred. No. 3.6e-36;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 176 AQPGLVFFADDDNTYSLELFQEMRTTRKVSVPVGLVGGRRYERPLV 222
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 177 AQPGLVFFADDDNTYSLELFQEMRTTRKVSVPVGLVGGRRYERPLV 223

RESULT 8
US-10-634-905-4
; Sequence 4, Application US/10634905
; Publication No. US20040067225A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; FILE REFERENCE: CL001147-DIV
; CURRENT APPLICATION NUMBER: US/10/634,905
; CURRENT FILING DATE: 2003-08-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-634-905-4

Query Match      14.6%; Score 47; DB 4; Length 324;
Best Local Similarity 100.0%; Pred. No. 3.6e-36;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 176 AQPGLVFFADDDNTYSLELFQEMRTTRKVSVPVGLVGGRRYERPLV 222
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 177 AQPGLVFFADDDNTYSLELFQEMRTTRKVSVPVGLVGGRRYERPLV 223

RESULT 9
US-10-029-386-33145
; Sequence 33145, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33145
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```

; OTHER INFORMATION: MAP TO AC013591.3  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6  
; OTHER INFORMATION: SWISSPROT HIT: Q9P2W7, EVALUAE 9.00e-18  
US-10-029-386-33145

Query Match 4.0%; Score 13; DB 4; Length 40;  
Best Local Similarity 100.0%; Pred. No. 0.0002;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 FADDNTNYSLELF 195

DB 26 FADDNTNYSLELF 38

RESULT 10  
US-10-450-763-43912  
; Sequence 43912, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 43912  
; LENGTH: 167  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-450-763-43912

Query Match 3.7%; Score 12; DB 5; Length 167;  
Best Local Similarity 100.0%; Pred. No. 0.0062;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 238 RPPAIDMAGFAV 249

DB 42 RPPAIDMAGFAV 53

RESULT 11  
US-10-698-190-16  
; Sequence 16, Application US/10698190  
; Publication No. US20040147469A1  
; GENERAL INFORMATION:  
; APPLICANT: Silver, J.  
; APPLICANT: Grimpe, B.  
; TITLE OF INVENTION: Methods of Inhibiting Glial Scar Formation  
; FILE REFERENCE: CWRU-P01-018  
; CURRENT APPLICATION NUMBER: US/10/698,190  
; CURRENT FILING DATE: 2003-10-31  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 16  
; LENGTH: 335  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-698-190-16

Query Match 3.4%; Score 11; DB 4; Length 335;  
Best Local Similarity 100.0%; Pred. No. 0.1;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 204 VSWVPVGLVGG 214

DB 213 VSWVPVGLVGG 223

RESULT 12

US-10-723-860-4238  
; Sequence 4238, Application US/10723860  
; Publication No. US20040253606A1  
; GENERAL INFORMATION:  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Ginsburg, Wendy M.  
; APPLICANT: Zlotnik, Albert  
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &  
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators  
; FILE REFERENCE: 05882.0193.NPUS01  
; CURRENT APPLICATION NUMBER: US/10/723,860  
; CURRENT FILING DATE: 2003-11-26  
; PRIOR APPLICATION NUMBER: 60/429,739  
; PRIOR FILING DATE: 2002-11-26  
; NUMBER OF SEQ ID NOS: 8393  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 4238  
; LENGTH: 335  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-723-860-4238

Query Match 3.4%; Score 11; DB 5; Length 335;  
Best Local Similarity 100.0%; Pred. No. 0.1;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 204 VSWVPVGLVGG 214

DB 213 VSWVPVGLVGG 223

RESULT 13

US-10-450-763-30691  
; Sequence 30691, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 30691  
; LENGTH: 691  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (583)..(632)  
; OTHER INFORMATION: PHOSPHORYLASE KINASE ALPHA MUSCL domain identified by  
; OTHER INFORMATION: eMATRIX, accession number PD01841F, p-value=7.048e-09, raw score  
; OTHER INFORMATION: 13.36  
US-10-450-763-30691

Query Match 3.4%; Score 11; DB 5; Length 691;  
Best Local Similarity 100.0%; Pred. No. 0.19;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 204 VSWVPVGLVGG 214

DB 348 VSWVPVGLVGG 358

```
RESULT 14
US-10-403-571-34
; Sequence 34, Application US/10403571
; Publication No. US20040068763A1
; GENERAL INFORMATION:
; APPLICANT: Hopkins, Nancy
; APPLICANT: Golling, Gregory
; APPLICANT: Amsterdam, Adam
; APPLICANT: Sun, Zhaoxia
; TITLE OF INVENTION: Developmental Mutations in Zebrafish
; FILE REFERENCE: 01997/539002
; CURRENT APPLICATION NUMBER: US/10/403,571
; CURRENT FILING DATE: 2003-03-25
; PRIOR APPLICATION NUMBER: US 60/368,760
; PRIOR FILING DATE: 2002-03-29
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Danio rerio
US-10-403-571-34

Query Match      3.1%; Score 10; DB 4; Length 201;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 183 FADDNTYSL 192
Db 187 FADDNTYSL 196

RESULT 15
US-10-425-115-346410
; Sequence 346410, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 346410
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_79090C.1.pep
US-10-425-115-346410

Query Match      2.5%; Score 8; DB 4; Length 57;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 133 AGLPSTHL 140
Db 47 AGLPSTHL 54

Search completed: May 9, 2006, 11:00:09
Job time : 79 secs
```

**This Page Blank (uspto)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2006, 10:59:55 ; Search time 12 Seconds  
(without alignments)  
1245.831 Million cell updates/sec

Title: US-10-634-905-2

Perfect score: 323

Sequence: 1 MKSALFRFFILLPWILVI.....EKVLANEPKYHLDTVKIEV 323

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 235405 seqs, 46284737 residues

Word size : 1

Total number of hits satisfying chosen parameters: 235171

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published Applications AA New:\*

- 1: /SID55/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /SID55/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 3: /SID55/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 4: /SID55/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 5: /SID55/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 6: /SID55/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 7: /SID55/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 8: /SID55/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 9: /SID55/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 10: /SID55/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 11: /SID55/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 12: /SID55/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	323	100.0	323	9	US-10-330-773-566
2	53	16.4	290	9	US-10-330-773-563
3	8	2.5	498	9	US-10-467-657-5006
4	7	2.2	107	11	US-11-264-096-871
5	7	2.2	144	11	US-11-096-568A-26178
6	7	2.2	160	11	US-11-096-568A-26177
7	7	2.2	199	11	US-11-079-463-9881
8	7	2.2	205	11	US-11-113-424-42
9	7	2.2	212	11	US-11-188-298-12238
10	7	2.2	237	11	US-11-188-298-3599
11	7	2.2	237	11	US-11-188-298-4760
12	7	2.2	237	11	US-11-188-298-12773
13	7	2.2	258	11	US-11-096-568A-182
14	7	2.2	263	11	US-11-072-512-2192
15	7	2.2	293	11	US-11-096-568A-16425
16	7	2.2	295	11	US-11-096-568A-30519
17	7	2.2	319	11	US-11-087-099-2924
18	7	2.2	329	11	US-11-096-568A-30518
19	7	2.2	332	11	US-11-188-298-18650
20	7	2.2	358	9	US-10-793-626-18
21	7	2.2	358	11	US-11-188-298-14936

22	7	2.2	387	11	US-11-098-686-11142	Sequence 11142, A
23	7	2.2	408	11	US-11-140-417-6	Sequence 6, Appl
24	7	2.2	428	11	US-11-229-484-16	Sequence 16, Appl
25	7	2.2	472	11	US-11-087-099-10910	Sequence 10910, A
26	7	2.2	472	11	US-11-188-298-21080	Sequence 21080, A
27	7	2.2	488	11	US-11-079-463-6537	Sequence 6537, Ap
28	7	2.2	569	11	US-11-096-568A-30517	Sequence 30517, A
29	7	2.2	600	11	US-11-079-463-9991	Sequence 9991, Ap
30	7	2.2	842	11	US-11-188-298-6227	Sequence 6227, Ap
31	7	2.2	1574	9	US-10-055-877-211	Sequence 211, App
32	6	1.9	17	11	US-11-202-449-3	Sequence 3, Appli
33	6	1.9	33	9	US-10-467-657-8872	Sequence 8872, Ap
34	6	1.9	34	11	US-11-207-078-318	Sequence 318, App
35	6	1.9	39	11	US-11-088-785-24	Sequence 24, Appl
36	6	1.9	42	11	US-11-088-785-23	Sequence 23, Appl
37	6	1.9	52	11	US-11-088-785-21	Sequence 21, Appl
38	6	1.9	54	11	US-11-088-785-22	Sequence 22, Appl
39	6	1.9	71	9	US-10-467-657-6652	Sequence 6652, Ap
40	6	1.9	72	9	US-10-508-307-7	Sequence 7, Appli
41	6	1.9	76	9	US-10-821-234-910	Sequence 910, App
42	6	1.9	83	9	US-10-510-386-114	Sequence 114, App
43	6	1.9	83	9	US-10-895-064-2389	Sequence 2389, Ap
44	6	1.9	83	11	US-11-129-741-2389	Sequence 2389, Ap
45	6	1.9	96	9	US-10-475-075-809	Sequence 809, App

#### ALIGNMENTS

RESULT 1

US-10-330-773-566

; Sequence 566, Application US/10330773

; Publication No. US20060040262A1

; GENERAL INFORMATION:

; APPLICANT: David W. Morris

; APPLICANT: Marc Malandro

; TITLE OF INVENTION: Novel Compositions and Methods in Cancer

; FILE REFERENCE: 529452001300

; CURRENT APPLICATION NUMBER: US/10/330,773

; CURRENT FILING DATE: 2002-12-27

; NUMBER OF SEQ ID NOS: 981

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 566

; LENGTH: 323

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-330-773-566

Query Match 100.0%; Score 323; DB 9; Length 323;

Best Local Similarity 100.0%; Pred. No. 2.2e-317;

Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MKSALFRFFILLPWILVIIMLDVDTTRRPVPLTPRPFPSPYAVGEGGARLPLRRGCPA	60
Db	1	MKSALFRFFILLPWILVIIMLDVDTTRRPVPLTPRPFPSPYAVGEGGARLPLRRGCPA	60
Qy	61	HGTQKNQSRPQPPQLPTTIVAITPTYSRPVQKAELELANTERQVQLHLLVEDAA	120
Db	61	HGTQKNQSRPQPPQLPTTIVAITPTYSRPVQKAELELANTERQVQLHLLVEDAA	120
Qy	121	ARSELVSRFLARAGLSTHLHVPTRPRYKRPGLPRATEQRNAGLWLRQRHQHQAQPGV	180
Db	121	ARSELVSRFLARAGLSTHLHVPTRPRYKRPGLPRATEQRNAGLWLRQRHQHQAQPGV	180
Qy	181	LPFADDNTYSLFELFEMRTTRKVSVPVGLVGGRYERPLVNGKVGVYTGWRADRPF	240
Db	181	LPFADDNTYSLFELFEMRTTRKVSVPVGLVGGRYERPLVNGKVGVYTGWRADRPF	240
Qy	241	AIDMAGFASLVQILSNPKAVFKRGSCQMQSSDFLKQITTVVELEPKANNCTKVLVWH	300
Db	241	AIDMAGFASLVQILSNPKAVFKRGSCQMQSSDFLKQITTVVELEPKANNCTKVLVWH	300
Qy	301	TRTEKVNLANEPKYHLDTVKIEV	323

```
Db      301 TRTEKVNLANEPKXHLDTWKIEV 323
|||||
RESULT 2
US-10-330-773-563
; Sequence 563, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330, 773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 563
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-330-773-563

Query Match      16.4%; Score 53; DB 9; Length 290;
Best Local Similarity 100.0%; Pred. No. 2.2e-45;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      237 DRPFAIDMAGFAVSLQVILSNKAVFKRRGSGPQMGSDFLKQITTVESLEPK 289
|||||
Db      206 DRPFAIDMAGFAVSLQVILSNKAVFKRRGSGPQMGSDFLKQITTVESLEPK 258
|||||

RESULT 3
US-10-467-657-5006
; Sequence 5006, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 5006
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-5006

Query Match      2.5%; Score 8; DB 9; Length 498;
Best Local Similarity 100.0%; Pred. No. 8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      222 VENGKVVVG 229
|||||
Db      136 VENGKVVVG 143
|||||

RESULT 4
US-11-264-096-871
; Sequence 871, Application US/11264096
; Publication No. US20060084794A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546D1
```

```
; CURRENT APPLICATION NUMBER: US/11/264,096
; CURRENT FILING DATE: 2005-11-02
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 871
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (66)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-11-264-096-871

Query Match      2.2%; Score 7; DB 11; Length 107;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      46 GRGGARL 52
|||||
Db      101 GRGGARL 107
|||||

RESULT 5
US-11-096-568A-26178
; Sequence 26178, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 26178
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(144)
; OTHER INFORMATION: Ceres Seq. ID no. 13498759
US-11-096-568A-26178

Query Match      2.2%; Score 7; DB 11; Length 144;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      53 PLRRGGP 59
|||||
Db      59 PLRRGGP 65
|||||

RESULT 6
US-11-096-568A-26177
; Sequence 26177, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
```

```
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 26177
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(160)
; OTHER INFORMATION: Ceres Seq. ID no. 13498758
US-11-096-568A-26177

Query Match          2.2%; Score 7; DB 11; Length 160;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 53 PLRRGGP 59
Db 75 PLRRGGP 81

RESULT 7
US-11-079-463-9881
; Sequence 9881, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FRATRICOLA
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 9881
; LENGTH: 199
; TYPE: PRT
; ORGANISM: B. fragilis
US-11-079-463-9881

Query Match          2.2%; Score 7; DB 11; Length 199;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 96 AELTRLA 102
Db 4 AELTRLA 10

RESULT 8
US-11-113-424-42
; Sequence 42, Application US/11113424
; Publication No. US20050260713A1
; GENERAL INFORMATION:
; APPLICANT: Gargolli et al.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-225
; CURRENT APPLICATION NUMBER: US/11/113,424
; CURRENT FILING DATE: 2005-04-21
; PRIOR APPLICATION NUMBER: 60/256,704
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/311,590
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/257,314
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/311,613
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/315,617
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/307,506
; PRIOR FILING DATE: 2001-07-24

; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 26177
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(160)
; OTHER INFORMATION: Ceres Seq. ID no. 13498758
US-11-096-568A-26177

; PRIOR APPLICATION NUMBER: 60/322,358
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/294,075
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/288,153
; PRIOR FILING DATE: 2001-05-02
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-113-424-42

Query Match          2.2%; Score 7; DB 11; Length 205;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 51 RLPLRRG 57
Db 152 RLPLRRG 158

RESULT 9
US-11-188-298-12238
; Sequence 12238, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 12238
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Paenibacillus macerans
US-11-188-298-12238

Query Match          2.2%; Score 7; DB 11; Length 212;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 GPAHGQTQ 64
Db 94 GPAHGQTQ 100

RESULT 10
US-11-188-298-3589
; Sequence 3589, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 3589
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Paenibacillus macerans
US-11-188-298-3589

Query Match          2.2%; Score 7; DB 11; Length 237;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 182
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(258)
; OTHER INFORMATION: Ceres Seq. ID no. 14305256
US-11-096-568A-182

Query Match      2.2%; Score 7; DB 11; Length 258;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 VGLVGR 215
Db 140 VGLVGR 146

RESULT 14
US-11-072-512-2192
; Sequence 2192, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2192
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-2192

Query Match      2.2%; Score 7; DB 11; Length 263;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 LVGGRRY 217
Db 189 LVGGRRY 195

RESULT 15
US-11-096-568A-16425
; Sequence 16425, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PGS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
```

```
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 182
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(258)
; OTHER INFORMATION: Ceres Seq. ID no. 14305256
US-11-096-568A-182

Query Match      2.2%; Score 7; DB 11; Length 258;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 VGLVGR 215
Db 140 VGLVGR 146

RESULT 14
US-11-072-512-2192
; Sequence 2192, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2192
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-2192

Query Match      2.2%; Score 7; DB 11; Length 263;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 LVGGRRY 217
Db 189 LVGGRRY 195

RESULT 15
US-11-096-568A-16425
; Sequence 16425, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PGS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
```

```
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 182
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(258)
; OTHER INFORMATION: Ceres Seq. ID no. 14305256
US-11-096-568A-182

Query Match      2.2%; Score 7; DB 11; Length 258;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 VGLVGR 215
Db 140 VGLVGR 146

RESULT 14
US-11-072-512-2192
; Sequence 2192, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2192
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-2192

Query Match      2.2%; Score 7; DB 11; Length 263;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 LVGGRRY 217
Db 189 LVGGRRY 195

RESULT 15
US-11-096-568A-16425
; Sequence 16425, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PGS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
```

; APPLICANT: Alexandrov, Nickolai et al.  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; FILE OF INVENTION: Therby  
; FILE REFERENCE: 2750-1592PUS2  
; CURRENT APPLICATION NUMBER: US/11/096,568A  
; CURRENT FILING DATE: 2005-04-01  
; NUMBER OF SEQ ID NOS: 34471  
; SEQ ID NO 16425  
; LENGTH: 293  
; TYPE: PRT  
; ORGANISM: Zea mays subsp. mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(293)  
; OTHER INFORMATION: Ceres Seq. ID no. 12352233  
US-11-096-568A-16425

Query Match 2.2%; Score 7; DB 11; Length 293;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 VGRGGAR 51  
|||  
Db 69 VGRGGAR 75

Search completed: May 9, 2006, 11:00:26  
Job time : 12 secs

**This Page Blank (uspto)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2006, 10:51:59 ; Search time 110 Seconds  
(without alignments)  
2071.688 Million cell updates/sec

Title: US-10-634-905-2

Perfect score: 1705

Sequence: 1 MKSALFTRFPILLWILVI.....EKVNLANEPKYHLDTVKIEV 323

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt\_05\_80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1705	100.0	323	1 B3GA2_HUMAN	Q9npz5 homo sapien
2	1705	100.0	323	2 Q599K0_PANTR	Q599k0 pan troglod
3	1620	95.0	329	1 B3GA2_CANFA	O5caz6 canis famil
4	1592.5	93.4	326	2 Q599K1_BOVIN	Q599k1 bos taurus
5	1531.5	89.8	324	1 B3GA2_MOUSE	P59270 mus musculu
6	1531.5	89.5	324	2 Q5DTT7_MOUSE	Q5dtt7 mus musculu
7	1526.5	89.5	324	1 B3GA2_RAT	Q9z137 rattus norv
8	1221	71.6	331	2 Q599J8_XENTR	Q599j8 xenopus tro
9	1211.5	71.1	304	2 Q5CA57_CHICK	O5ca57 gallus gall
10	1194.5	70.1	330	2 Q599J7_XENLA	Q599j7 xenopus lae
11	1154.5	67.7	304	2 Q5CA27_FUGRU	O5ca27 fugu rubrip
12	1121.5	65.8	304	2 Q5CA28_TETNG	O5ca28 tetraodon n
13	1102.5	64.7	316	2 Q6DBR2_BRARE	Q6db2r brachydanio
14	1057.5	62.0	413	2 Q4SG61_TETNG	O4sg61 tetraodon n
15	1055	61.9	303	2 Q599J9_ORYLA	O599j9 oryzias lat
16	909.5	53.3	327	2 Q5CB07_BRARE	Q5cb07 brachydanio
17	787.5	46.2	334	2 Q96FS7_HUMAN	Q96fs7 homo sapien
18	787	46.2	335	2 Q5CA29_TETNG	O5ca29 tetraodon n
19	783.5	46.0	334	2 Q5CB02_BRARE	O5cb02 brachydanio
20	779.5	45.7	334	1 B3GA1_HUMAN	O9p2w7 homo sapien
21	779	45.7	335	2 Q5CB01_ORYLA	O5cb01 oryzias lat
22	777.5	45.6	334	1 B3GA1_RAT	Q35789 r galactosy
23	777.5	45.6	334	2 Q6PIG8_MOUSE	Q6pig8 mus musculu
24	777.5	45.6	347	2 Q8R531_MOUSE	Q8r531 mus musculu
25	777	45.6	335	2 Q5CB00_FUGRU	O5cb00 fugu rubrip
26	776	45.5	335	1 B3GA1_CANFA	O5cb03 c galactosy
27	776	45.5	359	2 Q4SEI4_TETNG	Q4sei4 tetraodon n
28	772	45.3	394	2 Q5DTF9_MOUSE	Q5dtf9 mus musculu
29	770.5	45.2	334	1 B3GA1_MOUSE	Q9cwt3 mus musculu
30	764	44.8	277	2 Q8RIV2_MOUSE	Q8riv2 mus musculu
31	760	44.6	345	2 Q8B1C8_MOUSE	Q8b1c8 mus musculu

Q599k2 bos taurus  
O5cb04 p galactosy  
Q7t180 gallus gall  
O5cb05 fugu rubrip  
O63zm6 xenopus lae  
Q480u6 tetraodon n  
Q5cb06 oryzias lat  
Q599k3 xenopus tro  
O567g3 brachydanio  
Q7t1e2 gallus gall  
Q7zzm8 xenopus lae  
Q8jhi9 brachydanio  
Q599k4 xenopus tro  
Q7t1e1 gallus gall

ALIGNMENTS

RESULT 1  
B3GA2\_HUMAN  
ID B3GA2\_HUMAN STANDARD; PRT; 323 AA.  
AC Q9NPZ5; Q5JS09; Q9TF38; Q96NK4;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DE Galactosylgalactosylxylosylprotein 3-beta-glucuronosyltransferase 2  
DE (EC 2.4.1.135) (Beta-1,3-glucuronosyltransferase 2)  
DE (Glucuronosyltransferase-S) (GlcAT-S) (UDP-glucuronosyltransferase-S)  
DE (GlcAT-D).  
GN Name=B3GAT2; Synonyms=GLCATS, KIAA1963;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22410182; PubMed=12522689; DOI=10.1007/s100380200103;  
RA Marcos I., Galan J.J., Borrego S., Antinolo G.;  
RT "Cloning, characterization, and chromosome mapping of the human GlcAT-S gene.";  
RT J. Hum. Genet. 47:677-680(2002).  
RL [2]  
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
RP TISSUE=Brain;  
RX MEDLINE=21842142; PubMed=11853319;  
RA Nagase T., Kikuno R., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XXII. The complete sequences of 50 new cDNA clones which code for large proteins.";  
RT DNA Res. 8:319-327(2001).  
RL [3]  
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RX MEDLINE=22935763; PubMed=14574404; DOI=10.1038/nature02055;  
RA Mungall A.J., Palmer S.A., Sims S.K., Edwards C.A., Ashurst J.L., Wilming L., Jones M.C., Horton R., Hunt S.E., Scott C.E., Gilbert J.G.R., Clamp M.E., Bethel G., Milne S., Ainscough R., Almeida J.P., Ambrose K.D., Andrews T.D., Ashwell R.I.S., Babage A.K., Baguley C.L., Bailey J., Banerjee R., Barker D.J., Barlow K.F., Bates K., Beare D.M., Beasley H., Beasley O., Bird C.P., Blakey S.E., Bray-Allen S., Brook J., Brown A.J., Brown J.Y., Burford D.C., Burrill W., Burton J., Carter C., Carter N.P., Chapman J.C., Clark S.Y., Clark G., Clee C.M., Clegg S., Cobley V., Cullier R.E., Collins J.E., Colman L.K., Corby N.R., Coville G.J., Cullley K.M., Dhali P., Davies J., Dunn M., Earthrowl M.E., Ellington A.E., Evans K.A., Faulkner L., Francis M.D., Frankish A., Frankland J., French L., Garner P., Garnett J., Ghori M.J., Gilby L.M., Gillson C.J., Glithero R.J., Graham D.V., Grant M., Gribble S., Griffiths C., Griffiths M.N.D., Hall R., Halls K.S., Hammond S., Harley J.D., Hart E.A., Heath P.D., Heathcote R., Holmes S.J., Howden P.J., Howe K.L., Howell G.R., Huckle E., Humphray S.J., Humphries M.D., Hunt A.R., Johnson C.M., Joy A.A.,

RA Kay M., Keenan S.J., Kimberley A.M., King A., Laird G.K., Langford C.,  
RA Lawlor S., Leongamornlert D.A., Leversha M., Lloyd C.R., Lloyd D.M.,  
RA Loveland J.E., Lovell J., Martin S., Mashreghi-Sohamadi M.,  
RA Maslen G.L., Matthews L., McCann O.T., McLaren S.J., McWay K.,  
RA McMurray A., Moore M.J.F., Mullikin J.C., Niblett D., Nickerson T.,  
RA Novik K.L., Oliver K., Overton-Larty E.K., Parker A., Patel R.,  
RA Pearce A.V., Peck A.I., Phillimore B.J.C.T., Phillips S., Plumb R.W.,  
RA Porter K.M., Ramsey Y., Ranby S.A., Rice C.M., Ross M.T., Searle S.M.,  
RA Sehra H.K., Sheridan E., Skuce C.D., Smith S., Smith M., Spraggon L.,  
RA Squares S.L., Steward C.A., Sycamore N., Tamlyn-Hall G., Tester J.,  
RA Theaker A.J., Thomas D.W., Thorpe A., Tracey A., Tromans A., Tubby B.,  
RA Wall M., Wallis J.M., West A.P., White S.S., Whitehead S.L.,  
RA Whittaker H., Wild A., Willey D.J., Wilmer T.E., Wood J.M., Wray P.W.,  
RA Wyatt J.C., Young L., Younger R.M., Bentley D.R., Coulson A.,  
RA Durbin R., Hubbard T., Sulston J.E., Dunham I., Rogers J., Beck S.;  
RT "The DNA sequence and analysis of human chromosome 6.";  
RL Nature 425:805-811(2003).  
RN [4]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 157-323.  
RC TISSUE=fetal brain;  
RX PubMed=14702039; DOI=10.1038/ng1285;  
RA Oka T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,  
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,  
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Iehi S.,  
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,  
RA Nagahara K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,  
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,  
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,  
RA Omura Y., Abe K., Kamihara K., Katsuma T., Sato K., Tanikawa M.,  
RA Yamazaki M., Ninomiya K., Iehibashi T., Yamashita H., Murakawa K.,  
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,  
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hotuta T.,  
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,  
RA Nomura Y., Togiya S., Konai F., Hara R., Takeuchi K., Arita M.,  
RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,  
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara T., Sano S.,  
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,  
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,  
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,  
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,  
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,  
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,  
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
RA Okitani R., Kawakami T., Noguchi H., Itoh T., Shigeta K., Senba T.,  
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,  
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,  
RA Mizushima-Sugano J., Sato T., Shirai Y., Takahashi Y., Nakagawa K.,  
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,  
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;  
RT "Complete sequencing and characterization of 21,243 full-length human  
cDNAs.";  
RL Nat. Genet. 36:40-45(2004).  
CC -!- FUNCTION: Involved in the biosynthesis of L2/HNK-1 carbohydrate  
CC epitope on both glycolipids and glycoproteins (By similarity).  
CC -!- CATALYTIC ACTIVITY: UDP-glucuronate + 3-beta-D-galactosyl-4-beta-  
CC D-galactosyl-O-beta-D-xylosylprotein = UDP + 3-beta-D-  
CC glucuronosyl-3-beta-D-galactosyl-4-beta-D-galactosyl-O-beta-D-  
CC xylosylprotein.  
CC -!- COFACTOR: Manganese.  
CC -!- PATHWAY: Glycosylation.  
CC -!- SUBUNIT: Homodimer (Potential).  
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Golgi.  
CC -!- TISSUE SPECIFICITY: Expressed in the trachea, retina, spinal cord,  
CC hippocampus and other brain regions, and, at lower levels, in  
CC testis and ovary.  
CC -!- SIMILARITY: Belongs to the glycosyltransferase 43 family.  
CC  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.

DR EMBL: AY070019; AAL57718.1; -; mRNA.  
DR EMBL: AY070110; AAL58977.1; -; Genomic DNA.  
DR EMBL: AY070108; AAL58977.1; JOINED; Genomic DNA.  
DR EMBL: AY070109; AAL58977.1; JOINED; Genomic DNA.  
DR EMBL: AB075843; BAB85549.1; -; ALT INIT; mRNA.  
DR EMBL: AL121961; CA142145.1; -; Genomic DNA.  
DR EMBL: AL450320; CA142145.1; JOINED; Genomic DNA.  
DR EMBL: AL450320; CA139582.1; -; Genomic DNA.  
DR EMBL: AL121961; CA139582.1; JOINED; Genomic DNA.  
DR EMBL: AK055248; BAB70889.1; -; ALT INIT; mRNA.  
DR HSSP: O94766; 1KWS.  
DR Ensembl: ENSG00000112309; Homo sapiens.  
DR HGNC: HGNC:922; B3GAT2.  
DR H-InvDB: HIX00006000; -.  
DR MIM: 607497; -.  
DR GO: GO:0016021; C:integral to membrane; ISS.  
DR GO: GO:0015018; F:galactosylgalactosylxylosylprotein 3-beta-g. .; ISS.  
DR GO: GO:0016757; F:transferase activity, transferring glycosyl. .; ISS.  
DR GO: GO:0016051; P:carbohydrate biosynthesis; ISS.  
DR InterPro: IPR005027; Glyco\_trans\_43.  
DR PANTHER: PTHR10896; Glyco\_trans\_43; 1.  
DR Pfam: PF03360; Glyco\_trans\_43; 1.  
DR Glycoprotein; Golgi stack; Manganese; Metal-binding; Multigene family;  
KW Signal-anchor; Transferase; Transmembrane.  
FT TOPO\_DOM 1 2 Cytoplasmic (Potential).  
FT TRANSMEM 3 23 Signal-anchor for type II membrane  
FT TOPO\_DOM 24 323 Lumenal (Potential).  
FT ACT\_SITE 273 273 Proton acceptor (By similarity).  
FT METAL 187 187 Manganese (By similarity).  
FT CARBOHYD 67 67 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 292 292 N-linked (GlcNAc...) (Potential).  
SQ SEQUENCE 323 AA; 36919 MW; 85058D52D2D28463 CRC64;  
Query Match 100.0%; Score 1705; DB 1; Length 323;  
Best Local Similarity 100.0%; Pred. No. 2.2e-127; Indels 0; Gaps 0;  
Matches 323; Conservative 0; Mismatches 0;  
QY 1 MKSALFTRFFILLPWILVIMLDVDRPRPPLTPRPYFSPYAVGCGARLPLRRGGPA 60  
DB 1 MKSALFTRFFILLPWILVIMLDVDRPRPPLTPRPYFSPYAVGCGARLPLRRGGPA 60  
QY 61 HGTQRNRSRQPOPEPQLPTIATPTYSRVPQKELTRLANTPROVAQLHWILVEDAA 120  
DB 61 HGTQRNRSRQPOPEPQLPTIATPTYSRVPQKELTRLANTPROVAQLHWILVEDAA 120  
QY 121 ARSELVSRFLARAGLPSTHLVPTPRRYKRPGLPRATEORNAGLAWLRHORHQRAQPGV 180  
DB 121 ARSELVSRFLARAGLPSTHLVPTPRRYKRPGLPRATEORNAGLAWLRHORHQRAQPGV 180  
QY 181 LFFADDNDNTYSLELFQEMRTTRKVSVPVGLVGGRRYERPLVENGKVGWYTGWRADRP 240  
DB 181 LFFADDNDNTYSLELFQEMRTTRKVSVPVGLVGGRRYERPLVENGKVGWYTGWRADRP 240  
QY 241 AIDMAGFAVSLQVILSNKPAVKFRGSGQPMQESDFLQKITTVVELEPKANNCTKVLVWH 300  
DB 241 AIDMAGFAVSLQVILSNKPAVKFRGSGQPMQESDFLQKITTVVELEPKANNCTKVLVWH 300  
QY 301 TRTEKVNLANEPKYHLDVTVKIEV 323  
DB 301 TRTEKVNLANEPKYHLDVTVKIEV 323  
RESULT 2  
Q599K0 PANTR PRELIMINARY; PRT; 323 AA.  
AC Q599K0;  
DT 10-MAY-2005 (TrEMBLrel. 30, Created)  
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)  
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)  
DE Beta-3-glucuronyltransferase-S (EC 2.4.1.135).  
GN Name=b3gat2;  
OS Pan troglodytes (Chimpanzee).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Pan.
OX NCBI_TaxID=9598;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ouzine M., Magdalou J., Fournel-Gigleux S., Mollicone R., Oriol R.;
RT "Phylogeny of beta3-glucuronosyltransferases.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ890841; CAI68027.1; -; mRNA.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 323 AA; 36918 MW; 85058D5D2D28463 CRC64;

Query Match 100.0%; Score 1705; DB 2; Length 323;
Best Local Similarity 100.0%; Pred. No. 2.2e-127;
Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKSALFRFFILLPWILLIIMLDVTRRPVPLTPRPYPSPYAVGGRGRLPLRRGGPA 60
Db 1 MKSALFRFFILLPWILLIIMLDVTRRPVPLTPRPYPSPYAVGGRGRLPLRRGGPA 60
Qy 61 HGTRKNSRQPOPEPOLPTIYAITPTYSRPVQKAELTRANTFROVAQLHWLVEDAA 120
Db 61 HGTRKNSRQPOPEPOLPTIYAITPTYSRPVQKAELTRANTFROVAQLHWLVEDAA 120
Qy 121 ARSELVSRFLARAGLPSTHLHVPTRRYKRPGLPRATEQRNAGLAWLRQRHQHQAQGV 180
Db 121 ARSELVSRFLARAGLPSTHLHVPTRRYKRPGLPRATEQRNAGLAWLRQRHQHQAQGV 180
Qy 181 LFFADDNTYSLELFQEMRTTRKYSVMPVGLVGRRYERPLVGVGWTGWDRDRPF 240
Db 181 LFFADDNTYSLELFQEMRTTRKYSVMPVGLVGRRYERPLVGVGWTGWDRDRPF 240
Qy 241 AIDMAGFAVSLQVLSNPKAVFKRGSGQPGQESDFLKQITTVBELSPKANCTKVLVWH 300
Db 241 AIDMAGFAVSLQVLSNPKAVFKRGSGQPGQESDFLKQITTVBELSPKANCTKVLVWH 300
Qy 301 TRTEKVNLANEPKYLHDTVKIEV 323
Db 301 TRTEKVNLANEPKYLHDTVKIEV 323

RESULT 3
B3GA2 CANFA
ID B3GA2 CANFA STANDARD; PRT; 329 AA.
AC Q5CAZ6;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Galactosylgalactosylxylosylprotein 3-beta-glucuronosyltransferase 2
DE (EC 2.4.1.135) (Beta-1,3-glucuronosyltransferase 2)
DE (Glucuronosyltransferase-S) (GlcAT-S) (UDP-glucuronosyltransferase-S)
DE (GlcAT-D).
GN Name=B3GAT2;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
OC Canis.
OX NCBI_TaxID=9615;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RA Ouzine M., Magdalou J., Fournel-Gigleux S., Mollicone R., Oriol R.;
RT "Phylogeny of beta3-glucuronosyltransferases.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Involved in the biosynthesis of L2/HNK-1 carbohydrate epitope on both glycolipids and glycoproteins (By similarity).
CC -1- CATALYTIC ACTIVITY: UDP-glucuronate + 3-beta-D-galactosyl-4-beta-D-galactosyl-1-O-beta-D-xylosylprotein = UDP + 3-beta-D-glucuronosyl-1-O-beta-D-xylosylprotein.
CC -1- COFACTOR: Manganese (By similarity).
CC -1- PATHWAY: Glycosylation.
CC -1- SUBUNIT: Homodimer (Potential).

CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (By similarity).
CC -1- SIMILARITY: Belongs to the glycosyltransferase 43 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
CC -----
DR EMBL; AJ888982; CAI62047.1; -; mRNA.
DR InterPro; IPR005027; Glyco_trans_43.
DR PANTHER; PTHR10896; Glyco_trans_43; 1.
DR Pfam; PF03360; Glyco_transf_43; 1.
KW Glycoprotein; Golgi stack; Manganese; Metal-binding; Multigene family; Signal-anchor; Transferase; Transmembrane.
FT TOPO_DOM 1 2 Cytoplasmic (Potential).
FT TRANSMEM 3 23 Signal-anchor for type II membrane protein (Potential).
FT TOPO_DOM 24 329 Lumenal (Potential).
FT ACT_SITE 279 279 Proton acceptor (By similarity).
FT MSTAL 193 193 Manganese (By similarity).
FT CARBOHYD 71 71 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 298 298 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 329 AA; 37561 MW; 45DF20A242CD1E72 CRC64;

Query Match 95.0%; Score 1620; DB 1; Length 329;
Best Local Similarity 94.5%; Pred. No. 1.3e-120;
Matches 311; Conservative 5; Mismatches 7; Indels 6; Gaps 2;

Qy 1 MKSALFRFFILLPWILLIIMLDVTRRPVPLTPRPYPSPYAVGGRGRLPLRRG--- 57
Db 1 MKSALFRFFILLPWILLIIMLDVTRRPVPLTPRPYPSPYAVGGRGRLPLRRGGPD 60
Qy 58 -GPAHTQKNSRQPOPEPOLPTIYAITPTYSRPVQKAELTRANTFROVAQLHWI 114
Db 61 SGPRGWEKESRPHARPEPLPTIYAITPTYSRPVQKAELTRANTFROVAQLHWI 120
Qy 115 LVEDAARSELVSRFLARAGLPSTHLHVPTRRYKRPGLPRATEQRNAGLAWLRQRHQH 174
Db 121 LVEDAARSELVSRFLARAGLPSTHLHVPTRRYKRPGLPRATEQRNAGLAWLRQRHQH 180
Qy 175 RAQPGVLFADDDNTYSLELFQEMRTTRKYSVMPVGLVGRRYERPLVGVGWTGW 234
Db 181 RAQPGVLFADDDNTYSLELFQEMRTTRKYSVMPVGLVGRRYERPLVGVGWTGW 240
Qy 235 RADRPFAIDMAGFAVSLQVLSNPKAVFKRGSGQPGQESDFLKQITTVBELSPKANCT 294
Db 241 RADRPFAIDMAGFAVSLQVLSNPKAVFKRGSGQPGQESDFLKQITTVBELSPKANCT 300
Qy 295 KVLVWHTRTEKVNLANEPKYLHDTVKIEV 323
Db 301 KVLVWHTRTEKVNLANEPKYLHDTVKIEV 329

RESULT 4
Q599K1 BOVIN
ID Q599K1 BOVIN PRELIMINARY; PRT; 326 AA.
AC Q599K1;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Beta-3-glucuronosyltransferase-S (EC 2.4.1.135).
GN Name=B3GAT2;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ouzine M., Magdalou J., Fournel-Gigleux S., Mollicone R., Oriol R.;
RT "Phylogeny of beta3-glucuronosyltransferases.";
```



```
DR EMBL; AB055902; BAC20343.1; -; mRNA.
DR EMBL; AK048146; BAC33257.1; -; mRNA.
DR EMBL; AK052640; BAC35075.1; -; mRNA.
DR EMBL; BC056368; AAH56368.1; -; mRNA.
DR EMBL; BC058082; AAH58082.1; -; mRNA.
DR HSSP; O94766; 1KWS.
DR Ensemble; ENSMUSG00000026156; Mus musculus.
DR MGI; MGI:2389490; B3gat2.
DR GO; GO:0016021; C: integral to membrane; ISS.
DR GO; GO:0015018; F: galactosylgalactosylxylosylprotein 3-beta-g. . .; ISS.
DR GO; GO:0016757; F: transferase activity, transferring glycosyl. . .; ISS.
DR GO; GO:0016031; F: carbohydrate biosynthesis; ISS.
DR InterPro; IPR005027; Glyco trans 43.
DR PANTHER; PTHR10896; Glyco trans 43; 1.
DR Pfam; PF03360; Glyco trans 43; 1.
KW Alternative splicing; Glycoprotein; Golgi stack; Manganese;
KW Metal-binding; Multigene family; Signal-anchor; Transferrase;
KW Transmembrane.
FT TOPO_DOM 1 2 Cytoplasmic (Potential).
FT TRANSMEM 3 23 Signal-anchor for type II membrane
FT FT protein (Potential).
FT FT Lumenal (Potential).
FT FT Proton acceptor (By similarity).
FT ACT_SITE 24 324 Manganese (by similarity).
FT METAL 274 274 N-linked (GlcNAc. . .) (potential).
FT CARBOHYD 68 188 N-linked (GlcNAc. . .) (potential).
FT CARBOHYD 293 293 FAVSLQVILSNPKAVFKRRGSGPQGMQSDFLKQITTVVEELE
FT VARSPPLIC 248 302 KPANCTKVLVWHT -> EONAWDIPCRRMGSPRENWERPIF
FT FT IKPLSVLHSSSEILLIFKLGSLMVQVEVDFRSG (in
FT FT isoform 2).
FT FT /FTID=VSP_001796.
FT FT Missing (in isoform 2).
FT FT /FTID=VSP_001797.
FT FT A -> V (in Ref. 1).
FT FT SEQUENCE 324 AA; 37132 MW; 675C549D3787B957 CRC64;
Query Match 89.8%; Score 1531.5; DB 1; Length 324;
Best Local Similarity 90.4%; Pred. No. 1.4e-113;
Matches 293; Conservative 10; Mismatches 20; Indels 1; Gaps 1;
Qy 1 MKSALFTRFFILLPWILVIIMLDVTRRPVPPPTPRPYFSPYAVGGRGGLPLRRGGPA 60
Db 1 MKSALCSRFLLPWILVIIMLDVTRRPVPPPTPRPYFSPYAVGGRGGLPLRRSSPG 60
Qy 61 H-GTQKNQSRPOPEPOLPTIYAITPTYSRPVQKAEELTLANTFRQVLAQLHWILVEDA 119
Db 61 RDAEKKNESRPQLQPEPLPTIYAITPTYSRPVQKAEELTLANTFRQVLAQLHWILVEDR 120
Qy 120 AARSELYSRFLARAGLPSTHLHVPTPRRYKRPGLPRATEQRNAGLAWLRQRHQHQAQPG 179
Db 121 ATRSELVSSFLARAGLPNTHLVPTPRRYKRPWLPATEQRNAGLAWLRQRHQHQAQPG 180
Qy 180 VLFPAADDNTYSLELFQEMRTTRKVSVPVGLVGGRRYRPLVNGKVGWYTTGWADRRP 239
Db 181 VLFPAADDNTYSLELFQEMRTTRKVSVPVGLVGGRRYRPLVNGKVGWYTTGWEDRRP 240
Qy 240 FAIDMAGFAVSLQVILSNPKAVFKRRGSGPQGMQSDFLKQITTVVEELEPKANCTKVLVW 299
Db 241 FAIDMAGFAVSLQVILSNPKAVFKRRGSGPQGMQSDFLKQITTVVEELEPKANCTKVLVW 300
Qy 300 HTRTEKVNLANEPKYHLDTVTKIEV 323
Db 301 HTRTEKVNLANEPKYHLDTVNIEV 324
RESULT 6
Q5D7T7 MOUSE PRELIMINARY; PRT; 349 AA.
AC Q5D7T7,
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE MKIAA1963 protein (Fragment).
Name=B3gat2; Synonyms=MKIAA1963;
Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
NUCLEOTIDE SEQUENCE.
TISSUE=Fetal brain;
Okazaki N., Kikuno R.F., Ohara R., Inamoto S., Nagase T., Ohara O.,
Koga H.,
"Prediction of the Coding Sequences of Mouse Homologues of KIAA Gene.
The Complete Nucleotide Sequences of Mouse KIAA-homologous cDNAs
Identified by Screening of Terminal sequences of cDNA Clones Randomly
Sampled from Size-Fractionated Libraries. ";
Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
RL EMBL; AK220433--BAD90265.1; -; mRNA.
DR MGI; MGI:2389490; B3gat2.
DR GO; GO:0015020; F: glucuronosyltransferase activity; TAS.
DR InterPro; IPR005027; Glyco trans 43.
DR Pfam; PF03360; Glyco trans 43; 1.
DR NON_TER 1
FT NON_TER 1
SQ SEQUENCE 349 AA; 40066 MW; 2EFB2B8BCBD84169 CRC64;
Query Match 89.8%; Score 1531.5; DB 2; Length 349;
Best Local Similarity 90.4%; Pred. No. 1.6e-113;
Matches 293; Conservative 10; Mismatches 20; Indels 1; Gaps 1;
Qy 1 MKSALFTRFFILLPWILVIIMLDVTRRPVPPPTPRPYFSPYAVGGRGGLPLRRGGPA 60
Db 26 MKSALCSRFLLPWILVIIMLDVTRRPVPPPTPRPYFSPYAVGGRGGLPLRRSSPG 85
Qy 61 H-GTQKNQSRPOPEPOLPTIYAITPTYSRPVQKAEELTLANTFRQVLAQLHWILVEDA 119
Db 86 RDAEKKNESRPQLQPEPLPTIYAITPTYSRPVQKAEELTLANTFRQVLAQLHWILVEDR 145
Qy 120 AARSELYSRFLARAGLPSTHLHVPTPRRYKRPGLPRATEQRNAGLAWLRQRHQHQAQPG 179
Db 146 ATRSELVSSFLARAGLPNTHLVPTPRRYKRPWLPATEQRNAGLAWLRQRHQHQAQPG 205
Qy 180 VLFPAADDNTYSLELFQEMRTTRKVSVPVGLVGGRRYRPLVNGKVGWYTTGWADRRP 239
Db 206 VLFPAADDNTYSLELFQEMRTTRKVSVPVGLVGGRRYRPLVNGKVGWYTTGWEDRRP 265
Qy 240 FAIDMAGFAVSLQVILSNPKAVFKRRGSGPQGMQSDFLKQITTVVEELEPKANCTKVLVW 299
Db 266 FAIDMAGFAVSLQVILSNPKAVFKRRGSGPQGMQSDFLKQITTVVEELEPKANCTKVLVW 325
RESULT 7
B3GA2 RAT
ID B3GA2 RAT STANDARD; PRT; 324 AA.
AC Q9ZL37,
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Galactosylgalactosylxylosylprotein 3-beta-glucuronosyltransferase 2
DE (EC 2.4.1.135) (Beta-1,3-glucuronyltransferase 2)
DE (Glucuronosyltransferase-S) (GLCAT-S) (UDP-glucuronosyltransferase-S)
DE (GLCAT-D).
GN Name=B3gat2; Synonyms=Glcats;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
[1]
NUCLEOTIDE SEQUENCE.
TISSUE=Brain;
Name=B3gat2; Synonyms=MKIAA1963;
Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
NUCLEOTIDE SEQUENCE.
TISSUE=Fetal brain;
Okazaki N., Kikuno R.F., Ohara R., Inamoto S., Nagase T., Ohara O.,
Koga H.,
"Prediction of the Coding Sequences of Mouse Homologues of KIAA Gene.
The Complete Nucleotide Sequences of Mouse KIAA-homologous cDNAs
Identified by Screening of Terminal sequences of cDNA Clones Randomly
Sampled from Size-Fractionated Libraries. ";
Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
RL EMBL; AK220433--BAD90265.1; -; mRNA.
DR MGI; MGI:2389490; B3gat2.
DR GO; GO:0015020; F: glucuronosyltransferase activity; TAS.
DR InterPro; IPR005027; Glyco trans 43.
DR Pfam; PF03360; Glyco trans 43; 1.
DR NON_TER 1
FT NON_TER 1
SQ SEQUENCE 349 AA; 40066 MW; 2EFB2B8BCBD84169 CRC64;
Query Match 89.8%; Score 1531.5; DB 2; Length 349;
Best Local Similarity 90.4%; Pred. No. 1.6e-113;
Matches 293; Conservative 10; Mismatches 20; Indels 1; Gaps 1;
Qy 1 MKSALFTRFFILLPWILVIIMLDVTRRPVPPPTPRPYFSPYAVGGRGGLPLRRGGPA 60
Db 26 MKSALCSRFLLPWILVIIMLDVTRRPVPPPTPRPYFSPYAVGGRGGLPLRRSSPG 85
Qy 61 H-GTQKNQSRPOPEPOLPTIYAITPTYSRPVQKAEELTLANTFRQVLAQLHWILVEDA 119
Db 86 RDAEKKNESRPQLQPEPLPTIYAITPTYSRPVQKAEELTLANTFRQVLAQLHWILVEDR 145
Qy 120 AARSELYSRFLARAGLPSTHLHVPTPRRYKRPGLPRATEQRNAGLAWLRQRHQHQAQPG 179
Db 146 ATRSELVSSFLARAGLPNTHLVPTPRRYKRPWLPATEQRNAGLAWLRQRHQHQAQPG 205
Qy 180 VLFPAADDNTYSLELFQEMRTTRKVSVPVGLVGGRRYRPLVNGKVGWYTTGWADRRP 239
Db 206 VLFPAADDNTYSLELFQEMRTTRKVSVPVGLVGGRRYRPLVNGKVGWYTTGWEDRRP 265
Qy 240 FAIDMAGFAVSLQVILSNPKAVFKRRGSGPQGMQSDFLKQITTVVEELEPKANCTKVLVW 299
Db 266 FAIDMAGFAVSLQVILSNPKAVFKRRGSGPQGMQSDFLKQITTVVEELEPKANCTKVLVW 325
RESULT 7
B3GA2 RAT
ID B3GA2 RAT STANDARD; PRT; 324 AA.
AC Q9ZL37,
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Galactosylgalactosylxylosylprotein 3-beta-glucuronosyltransferase 2
DE (EC 2.4.1.135) (Beta-1,3-glucuronyltransferase 2)
DE (Glucuronosyltransferase-S) (GLCAT-S) (UDP-glucuronosyltransferase-S)
DE (GLCAT-D).
GN Name=B3gat2; Synonyms=Glcats;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
[1]
NUCLEOTIDE SEQUENCE.
TISSUE=Brain;
```

RX MEDLINE=99185317; PubMed=10082676; DOI=10.1006/bbrc.1999.0151;  
 RA Seiki T., Oka S., Terayama K., Imiya K., Kawasaki T.;  
 RT "Molecular cloning and expression of a second glucuronyltransferase  
 involved in the biosynthesis of the HNK-1 carbohydrate epitope.";  
 RL Biochem. Biophys. Res. Commun. 255:182-187(1999).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE, AND CHARACTERIZATION.  
 RC TISSUE=Embryonic brain;  
 RX MEDLINE=99287914; PubMed=10358066; DOI=10.1074/jbc.274.24.17115;  
 RA Shimoda Y., Tajima Y., Nagase T., Harii K., Osumi N., Sanai Y.;  
 RT "Cloning and expression of a novel galactoside beta1,3-  
 glucuronyltransferase involved in the biosynthesis of HNK-1 epitope.";  
 RL J. Biol. Chem. 274:17115-17122(1999).  
 CC -1- FUNCTION: Involved in the biosynthesis of L2/HNK-1 carbohydrate  
 CC epitope on both glycolipids and glycoproteins. Substrates include  
 CC asialo-orosomucoid (ASOR), paragloboside (lacto-N-  
 CC neotetraosylceramide), Gal-beta-1,4-GlcNAc-beta-1,3-Gal-beta-1,4-  
 CC Glc-pyridylamine and Gal-beta-1,3-GlcNAc-beta-1,3-Gal-beta-1,4-  
 CC Glc-pyridylamine.  
 CC -1- CATALYTIC ACTIVITY: UDP-glucuronate + 3-beta-D-galactosyl-4-beta-  
 CC D-galactosyl-O-beta-D-xylosylprotein = UDP + 3-beta-D-  
 CC glucuronosyl-3-beta-D-galactosyl-4-beta-D-galactosyl-O-beta-D-  
 CC xylosylprotein.  
 CC -1- COFACTOR: Manganese.  
 CC -1- PATHWAY: Glycosylation.  
 CC -1- SUBUNIT: Homodimer (Potential).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Golgi.  
 CC -1- TISSUE SPECIFICITY: Expressed in the cerebral cortex, cerebellum  
 CC and whole brain.  
 CC -1- SIMILARITY: Belongs to the glycosyltransferase 43 family.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 DR EMBL; AB010441; BAA75219.1; -; mRNA.  
 DR EMBL; AF106624; AAD29576.1; -; mRNA.  
 DR PIR; JG0163; JG0163.  
 DR HSP; O94766; LKWS.  
 DR GO; GO:0016021; C:integral to membrane; NAS.  
 DR GO; GO:0015018; F:galactosylgalactosylxylosylprotein 3-beta-g. . . ; IDA.  
 DR GO; GO:0016757; P:transferase activity, transferring glycosyl. . . ; IEP.  
 DR GO; GO:0016051; P:carbohydrate biosynthesis; IEP.  
 DR InterPro; IPR005027; Glyco trans 43.  
 DR PANTHER; PTHR10896; Glyco trans 43; 1.  
 DR Pfam; PF03360; Glyco transf 43; 1.  
 DR Glycoprotein; Golgi stack; Manganese; Metal-binding; Multigene family;  
 KW Signal-anchor; Transferase; Transmembrane.  
 FT TOPO\_DOM 1 2  
 FT TRANSMEM 3 23  
 FT TOPO\_DOM 24 324  
 FT ACT\_SITE 274 274  
 FT METAL 188 188  
 FT CARBOHYD 68 68  
 FT CARBOHYD 293 293  
 FT CARBOHYD 324 AA; 37220 MW; 6283C29535B5D00E CRC64;  
 SQ SEQUENCE 324 AA; 37220 MW; 6283C29535B5D00E CRC64;  
 Query Match 89.5%; Score 1526.5; DB 1; Length 324;  
 Best Local Similarity 89.8%; Pred. No. 3.6e-113;  
 Matches 291; Conservative 10; Mismatches 22; Indels 1; Gaps 1;  
 QY 1 MKSALFTRFPILLPWILVIMLDVTRRPVPLTPRPYFSPYAVGCGARLPLRRGGPA 60  
 Db 1 MKSALCNRFPIILLPWILVIMLDVTRRPVPLTPRPYFSPYAVGCGARLPLRRSSPG 60  
 QY 61 H-GTKRNSRPPQPEPQLPTTIAITPTYSRPVQKAELTRANTFRQAQLHWILVEDA 119  
 Db 61 RDAAEKRNESRPPQPEPQLPTTIAITPTYSRPVQKAELTRANTFRQAQLHWILVEDR 120  
 QY 120 AARSELVSRFLARAGLPSTHLHVPTRRYKRPGLPRATEQRNAGLAWLRQHQRAQPG 179

Db 121 ATRSELVSRFLARAGLPSTHLHVPTRRYKRPGLPRATEQRNAGLAWLRQHQRAQPG 180  
 QY 180 VLFFADDDNTYSLLEFQEMRTTRKYSVMPVGLVGGRRYERPLVGVGWTGWRADRP 239  
 Db 181 VLFFADDDNTYSLLEFQEMRTTRKYSVMPVGLVGGRRYERPLVGVGWTGWRADRP 240  
 QY 240 FAIDNAGFAVSLQVILSNPKAVFKRSGSQPGQMSDFLKQITTTVEELPKANNCTKVLVW 299  
 Db 241 FAIDNAGFAVSLQVILSNPKAVFKRSGSQPGQMSDFLKQITTTVEELPKANNCTKVLVW 300  
 QY 300 HTRTEKVNLANEPKHYLDVTKIEV 323  
 Db 301 HTRTEKVNLANEPKHYMDVTNIEV 324  
 RESULT 8  
 Q599J8 XENTR PRELIMINARY; PRT; 331 AA.  
 AC Q599J8; (TREMBlrel. 30, Created)  
 DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)  
 DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)  
 DE Beta-3-glucuronyltransferase-S (EC 2.4.1.135).  
 GN Name-b3gat2;  
 OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 OC Xenopodinae; Xenopus; Silurana.  
 OX NCBI\_TaxID=8364;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Ouzaine M., Magdalou J., Fournel-Gigleux S., Mollicone R., Oriol R.;  
 RT "Phylogeny of beta3-glucuronyltransferases.";  
 RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ890843; CAI68029.1; -; mRNA.  
 KW Glycosyltransferase; Transferase.  
 SQ SEQUENCE 331 AA; 37869 MW; 21F64C20B8F239D8 CRC64;  
 Query Match 71.6%; Score 1221; DB 2; Length 331;  
 Best Local Similarity 74.1%; Pred. No. 7.6e-89;  
 Matches 246; Conservative 26; Mismatches 50; Indels 10; Gaps 5;  
 QY 1 MKSALFTRFPILLPWILVIMLDVTRR-PVPLTPRPYFSPYAVGCGARLPLRRGQP 59  
 Db 1 MKSVFYSRFPILLPWILVIMLDVTRSDTRSPGSGRVPTARYPYWNRKVG-BAALMKPPP 59  
 QY 60 AH-GTKRNSRPPQPEPQLPTTIAITPTYSRPVQKAELTRANTFRQAQLHWILVED 118  
 Db 60 SDPSTHQELLSATQRKNETVPVIAITPTYSRPVQKAELTRANTFRQVPRHLWILVED 119  
 QY 119 AARSELVSRFLARAGLPSTHLHVPTRRYKRPGLPRATEQRNAGLAWLRQHQH---OR 175  
 Db 120 SVHTELVSRFLAGAGVTSTHLYVTPRRYKRTGLPRATEQRNAGLAWLRQYQRPGLRT 179  
 QY 176 AQP----GVLPFADDNTYSLLEFQEMRTTRKYSVMPVGLVGGRRYERPLVGVGWTG 231  
 Db 180 AQFQDPTGVVFPFADDNTYSLLEFQEMRTTQKYSVMPVGLVGGRRYERPVVGVGWTG 239  
 QY 232 TCWRADRPFAIDNAGFAVSLQVILSNPKAVFKRSGSQPGQMSDFLKQITTTVEELPKAN 291  
 Db 240 TCWRADRPFAIDNAGFAVSLQVILSNPKAVFKRSGSQPGQMSDFLKQITTKVNELEPKAN 299  
 QY 292 NCTKVLVWHTRTEKVNLANEPKHYLDVTKIEV 323  
 Db 300 NCTKVLVWHTRTEKVNLANEPKHPQDITKIEV 331  
 RESULT 9  
 Q5CAS7 CHICK  
 ID Q5CAS7 CHICK PRELIMINARY; PRT; 304 AA.  
 AC Q5CAS7;  
 DT 10-MAY-2005 (TREMBlrel. 30, Created)

```

DT 10-MAY-2005 (TReMBLrel. 30, Last sequence update)
DE 10-MAY-2005 (TReMBLrel. 30, Last annotation update)
DE Beta3-glucuronyltransferase-S (EC 2.4.1.135).
GN Name=B3GAT2;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ouzaine M., Magdalou J., Fournel-Gigleux S., Mollicone R., Oriol R.;
RT "Phylogeny of beta3-glucuronyltransferases.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ896606; CAI63864.1; -; mRNA.
DR InterPro; IPR005027; Glyco_trans_43.
DR PANTHER; PTHR10896; Glyco_trans_43; 6.
DR Pfam; PF03360; Glyco_transf_43; 1.
DR Glycosyltransferase; Transferase.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 304 AA; 34618 MW; 5C45DF677FEEA624 CRC64;

Query Match 71.1%; Score 1211.5; DB 2; Length 304;
Best Local Similarity 91.7%; Pred. No. 3.9e-88;
Matches 233; Conservative 5; Mismatches 15; Indels 1; Gaps 1;

Qy 71 PQPOPEPQPIYAITPTYSRPVQKAEI... 130
Db 1 PPRPNATLPYIAYITPTYSRPVTAEL... 110

Qy 131 ARAGLPSTLHVPTPRYKRGPLPRATE... 189
Db 111 AGAGLPTLHVPTPRYKRGPLPRATE... 170

Qy 190 YSELFOEMRTTRKVSVPVGLVGGRRY... 249
Db 171 YSELFHEMRTTRKVSVPVGLVGGRRY... 230

Qy 250 SLQVILSNPKAVFKRGSPQMGESDFLK... 309
Db 231 SLQVILSHPKAVFKRGSPQMGESDFLK... 290

Qy 310 NEPKYHLDTVKIEV 323
Db 291 NEPKYHLDTVNIEV 304

RESULT 10
Q599J7 XENLA PRELIMINARY; PRT; 330 AA.
AC Q599J7;
DT 10-MAY-2005 (TReMBLrel. 30, Created)
DT 10-MAY-2005 (TReMBLrel. 30, Last sequence update)
DE Beta-3-glucuronyltransferase-S (EC 2.4.1.135).
GN Name=b3gat2;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ouzaine M., Magdalou J., Fournel-Gigleux S., Mollicone R., Oriol R.;
RT "Phylogeny of beta3-glucuronyltransferases.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ890844; CAI68030.1; -; mRNA.
DR Glycosyltransferase; Transferase.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 330 AA; 37637 MW; 9B99EE00CACD4F60 CRC64;

Query Match 70.1%; Score 1194.5; DB 2; Length 330;
Best Local Similarity 72.3%; Pred. No. 9.7e-87;
Matches 243; Conservative 27; Mismatches 47; Indels 19; Gaps 7;

```

```

Qy 1 MKSALFTRFFILLPWILVIIMLDVDR... 57
Db 1 MKSVFYTRFFILLPWILVIIMLDVDR... 58

Qy 58 GPA--HGTQKRNQSRPQPEPOLPIYAI... 115
Db 59 DPSTEHNSAPHNRGK---NETVLPVIF... 114

Qy 116 VEDAAARSELVSRFLARAGLPSTLHVPT... 174
Db 115 VEDSVHPTLVSRLPLAGAGVKSSHLV... 174

Qy 175 ---RAQP-----GVLFDADDNTYSL... 227
Db 175 GIHSAQPHDLGSGVFFADDNTYSLF... 234

Qy 228 VGYTGWRRADRPFAIDMAGFAVSLQVIL... 287
Db 235 VSWYTGWRADRPFAIDMAGFAVSLQV... 294

Qy 288 PKANNCTKVLVHTRTEKVNLANEPKYL... 323
Db 295 PKANNSTKVLVHTRTEKVNLANEPKH... 330

RESULT 11
Q5CAZ7 FUGRU PRELIMINARY; PRT; 304 AA.
AC Q5CAZ7;
DT 10-MAY-2005 (TReMBLrel. 30, Created)
DT 10-MAY-2005 (TReMBLrel. 30, Last annotation update)
DE Beta3-glucuronyltransferase (EC 2.4.1.135).
GN Name=b3gat2;
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ouzaine M., Magdalou J., Fournel-Gigleux S., Mollicone R., Oriol R.;
RT "Phylogeny of beta3-glucuronyltransferases.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ888981; CAI62046.1; -; mRNA.
DR InterPro; IPR005027; Glyco_trans_43.
DR InterPro; IPR008919; Retrov_capsid_N.
DR InterPro; IPR011991; Wing_hlx_DNA_Bd.
DR PANTHER; PTHR10896; Glyco_trans_43; 6.
DR Pfam; PF03360; Glyco_transf_43; 1.
DR Glycosyltransferase; Transferase.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 304 AA; 34657 MW; 0BD5F8A06B88E4E6 CRC64;

Query Match 67.7%; Score 1154.5; DB 2; Length 304;
Best Local Similarity 70.3%; Pred. No. 1.3e-83;
Matches 227; Conservative 29; Mismatches 48; Indels 19; Gaps 4;

Qy 1 MKSALFTRFFILLPWILVIIMLDVDR... 60
Db 1 MKSVFYSRFFILLPWILVIIMLDVDR... 52

Qy 61 HGTQKRNQSRPQPEPOLPIYAITPTYSR... 120
Db 53 -GPGAQNS-----ALPVIYAITPT... 103

Qy 121 ARSELVSRFLARAGLPSTLHVPTPRYK... 180
Db 104 TRIDMTWTRFLARCGVPTLHVPTPR... 161

Qy 181 LFFADDNTYSLFQEMRTTRKVSVPVGL... 240
Db 162 VFFADDNTYSLFQEMRSTRGVSVPVGL... 221

```

```
QY 241 AIDMAGFAVSLQVILSNPKAVFKRGSPQGMQESDFLKQITTVBELEPKANNCTKVLVWH 300
Db 222 ATDMGFAVNLQVILSNPKAVFKRGSPQGMQESDFLKQITTKVTDLEPKANNCTQVLVWH 281
QY 301 TRTEKVNLANEPKHYLDTVKIEV 323
Db 282 TRTEKPHLANEPKHKDTTVIEV 304

RESULT 12
ID Q5CAZ8_TETNG PRELIMINARY; PRT; 304 AA.
AC Q5CAZ8_TETNG PRELIMINARY; PRT; 304 AA.
DT 10-MAY-2005 (T-EMBLrel. 30, Created)
DT 10-MAY-2005 (T-EMBLrel. 30, Last sequence update)
DE 10-MAY-2005 (T-EMBLrel. 30, Last annotation update)
DE Beta3-glucuronyltransferase (EC 2.4.1.135).
GN Nameb3gat2;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ouzzine M., Magdalon J., Fournel-Gigleux S., Mollicone R., Oriol R.;
RT "Phylogeny of beta3-glucuronyltransferases.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ388580; CA162045.1; -; mRNA.
DR InterPro; IPR02282; Cytochrome_c.R.
DR InterPro; IPR005027; Glyco_trans_43.
DR InterPro; IPR008919; Retrov_capsid_N.
DR InterPro; IPR011991; Wing_hlx_DNA_Bd.
DR PANTHER; PTHR10896; Glyco_trans_43; 6.
DR Pfam; PF03360; Glyco_transf_43; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 304 AA; 34513 MW; C6DBABADA1D9A698 CRC64;

Query Match 65.8%; Score 1121.5; DB 2; Length 304;
Best Local Similarity 68.2%; Pred. No. 5.7e-81;
Matches 223; Conservative 30; Mismatches 47; Indels 27; Gaps 5;

QY 1 MKSALTRFFILLPWILVIMLDVTRRPVPLTPRPYFSPVAVGGGARLPRLRGGP- 59
Db 1 MKSVFFSRFFILLPWILVIMLDVDSKRTIR-----APAAAGRAARQARGSGAPG 52
QY 60 ---AHGTQKRNOSRPOPEPOLPTIYAITPTYSRPVQKAELTRANTPROVAQLHWILV 116
Db 53 AAGAH-----NQ8-----ALPVIYAITPTYRPVQKAELTRLAHAFQVPFRHWILV 99
QY 117 EDAAARSELVSRFLARAGLPSTHLHVTPRRYKRPGLPRATEORNAGLAWLRQHORA 176
Db 100 EDSTTRDWTFLACGVPYTHLVFTPRFRKTCMPRATEGORNALAWLQR--RGR 157
QY 177 QPGVLFADDNTYSLELFOEMRTTRKVSVPVGLVGGRRYERPLVYENGKVGWYTGWRA 236
Db 158 DAGVWFADDNTYSLELFEEMSTRGVSWPVGVGGRSYERPLVSEKVGWYTGWRP 217
QY 237 DRPFATDMAGFAVSLQVILSNPKAVFKRGSPQGMQESDFLKQITTVBELEPKANNCTKV 296
Db 218 DRPFATDMAGFAVNLQVILSNPKAVFKRGSPQGMQESDFLKQITTKVTDLEPKANNCTRV 277
QY 297 LVWHTTRTEKVNLANEPKHYLDTVKIEV 323
Db 278 LVWHTTRTEKPHLANEPKHKDTTVIEV 304

RESULT 13
Q6DBR2 BRARE
ID Q6DBR2 BRARE PRELIMINARY; PRT; 316 AA.
AC Q6DBR2;
```

```
DT 25-OCT-2004 (T-EMBLrel. 28, Created)
DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
DE Zgc:91925.
GN Name=Zgc:91925;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Rutterfield Y.S.N., Krzywiecki M.I., Skalska U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RG NIH MGC Project;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC078400; AAH78400.1; -; mRNA.
DR ZFIN; ZDB-GENE-040801-191; zgc:91925.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0015018; F:galactosylgalactosylxylosylprotein 3-beta-g. .; IEA.
DR InterPro; IPR005027; Glyco_trans_43.
DR PANTHER; PTHR10896; Glyco_trans_43; 1.
DR Pfam; PF03360; Glyco_transf_43; 1.
SQ SEQUENCE 316 AA; 36273 MW; 32EC11D94043836B CRC64;

Query Match 64.7%; Score 1102.5; DB 2; Length 316;
Best Local Similarity 68.7%; Pred. No. 1.9e-79;
Matches 222; Conservative 34; Mismatches 60; Indels 7; Gaps 4;

QY 1 MKSALTRFFILLPWILVIMLDVTRRPVPLTPRPYFSPVAVGGGARLPRLRGGPA 60
Db 1 MKSIFTRFFILLPWILVIMLDVTRKLSVRNTASFVLS--RLGNVQORQVTRTTS 58
QY 61 HGTQKRNOSRPOPEPOLPTIYAITPTYSRPVQKAELTRANTPROVAQLHWILVEDAA 120
Db 59 SSSTTASSS---GRNATSLPVIYAITPTYSRAVQKAELTRANTPROQPFHWIVEDAN 115
QY 121 ARSELVSRFLARAGLPSTHLHVTPRRYKRPGLPRATEORNAGLAWLRQHORAQPGV 180
Db 116 SITELVSRFLARCGVRYTHLVNFTPRFRKTCMPRATEGORNALGMIR-GHRGSK-DKGV 173
QY 181 LFFADDNDNTYSLELFOEMRTTRKVSVPVGLVGGRRYERPLVYENGKVGWYTGWRADRF 240
Db 174 VFFADDNDNTYSLELFEEMSTRGVSWPVGLVGGRRYERPLVYENGKVGWYTGWRADRF 233
QY 241 AIDMAGFAVSLQVILSNPKAVFKRGSPQGMQESDFLKQITTVBELEPKANNCTKVLVWH 300
Db 234 AIDMAGFAVNLQVILSNPKAVFKRGSPQGMQESDFLKQITKVEDLEPKANNCTQVLVWH 293
QY 301 TRTEKVNLANEPKHYLDTVKIEV 323
```

Search completed: May 9, 2006, 10:55:21  
Job time : 112 secs

**This Page Blank (uspto)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2006, 10:51:05 ; Search time 86 Seconds  
(without alignments)  
1650.225 Million cell updates/sec

Title: US-10-634-905-2

Perfect score: 1705

Sequence: 1 MKSALFTRFFILLPWILIVI.....EKNLANEPKHYLDTVKIEV 323

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 43937871 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_21:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*
- 9: Geneseqp2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1705	100.0	323	5	ADI03930 Human enz
2	1705	100.0	323	8	ADQ97589 Human can
3	1526.5	89.5	324	5	ADI03932 Rat galac
4	1223	71.7	292	8	ADQ97586 Mouse can
5	1039	60.9	197	5	ADI03933 HumanUDP-
6	787.5	46.2	347	5	ABB97603 Novel hum
7	779.5	45.7	334	6	ABO07264 Human p53
8	777.5	45.6	347	2	RAY08214 Rat Glucu
9	777.5	45.6	347	2	AAW90057 Rat GicAT
10	653	38.3	335	6	ABO07265 Human p53
11	653	38.3	335	8	ADO31194 Human GT
12	653	38.3	335	8	ADQ21418 Human sof
13	645	37.8	335	9	ADY70386 Galactosy
14	608	35.7	305	8	ADI00832 Fruit fly
15	510	29.9	241	4	AAW93879 Human pol
16	510	29.9	241	8	ADL31964 Human pro
17	502.5	29.5	316	8	ADI00836 Fruit fly
18	502.5	29.5	479	4	ABB62172 Drosophil
19	487.5	28.6	366	8	ADI00834 Fruit fly
20	487.5	28.6	443	4	ABB60238 Drosophil
21	422	24.8	145	4	AAW99900 Human exc
22	422	24.8	145	4	AAW43699 Human bla
23	422	24.8	145	8	ADF71658 Human bla
24	406	23.8	691	4	ABG00332 Novel hum

25	324.5	19.0	201	8	ADO57267	Ado57267 Kidney de
26	305.5	17.9	308	3	AAG36689	Aag36689 Arabidops
27	305.5	17.9	317	3	AAG36688	Aag36688 Arabidops
28	272.5	16.0	241	3	AAG36690	Aag36690 Arabidops
29	262	15.4	59	3	AAB45114	Aab45114 Human sec
30	246	14.4	479	8	ADY13062	Ady13062 Plant ful
31	241.5	14.2	460	8	ADY06832	Ady06832 Plant ful
32	227	13.3	411	8	ADX91479	Adx91479 Plant ful
33	209.5	12.3	316	8	ADX90903	Adx90903 Plant ful
34	194.5	11.4	167	4	ABG13553	Abg13553 Novel hum
35	193.5	11.3	492	5	ABP99383	Abp99383 Arabidops
36	193.5	11.3	492	5	ABP99383	Abp99383 Arabidops
37	180	10.6	59	3	AAB45113	Aab45113 Herbicida
38	171	10.0	273	8	ADX87904	Adx87904 Plant ful
39	169	9.9	544	5	ABB93266	Abb93266 Herbicida
40	169	9.9	544	5	ABB93284	Abb93284 Herbicida
41	150	8.8	29	3	AAB45116	Aab45116 Human sec
42	141.5	8.3	378	8	ADX72765	Adx72765 Plant ful
43	139	8.2	40	8	ABO59511	AbO59511 Human gen
44	137.5	8.1	267	8	ADX91941	Adx91941 Plant ful
45	130.5	7.7	521	7	ABO69373	AbO69373 Pseudomon

ALIGNMENTS

RESULT 1

ADI03930

ID ADI03930 standard; protein; 323 AA.

XX ADI03930;

DT 22-APR-2004 (first entry)

DE Human enzyme protein.

KW Enzyme protein; drug screening; disease diagnosis; human; gene therapy;  
chromosome 6; enzyme; glucuronyltransferase.

OS Homo sapiens.

PN WO200268657-A2.

XX 06-SEP-2002.

PF 08-FEB-2002; 2002WO-US003623.

PR 26-FEB-2001; 2001US-0270871P.

PR 26-MAR-2001; 2001US-00816095.

XX (PEKE ) PE CORP NY.

PI Gan W, Yan C, Merkulov GV, Di Francesco V, Beasley EM;

DR WPI; 2002-713380/77.

DR N-PSDB; ADI03929, ADI03931.

PT New human enzyme proteins, useful for treating or diagnosing disorders  
associated with abnormal expression of the protein, in drug screening  
assays and pharmacogenomic analysis.

XX Claim 1; SEQ ID NO 2; 127pp; English.

CC The invention relates to a novel isolated enzyme protein and encoding  
polynucleotides. The protein shows a high degree of similarity to a  
glucuronyltransferase cloned from a rabbit brain cDNA library. The  
peptides and nucleic acid molecules are useful as models for the  
development of human therapeutic targets, aid in the identification of  
therapeutic proteins, and serve as targets for the development of human  
therapeutic agents. The peptide may be used in drug screening assays, in  
assays to determine the biological activity of the protein, to raise  
antibodies or to elicit another immune response, as a reagent in assays  
designed to quantitatively determine levels of the protein in biological

CC fluids, or as markers for tissues in which the corresponding protein is  
CC preferentially expressed. The human enzyme protein is also useful for  
CC diagnosing a disease, predisposition to a disease, or treating a disorder  
CC characterized by an absence of, inappropriate or unwanted expression of  
CC the protein. The antibodies are useful in pharmacogenomic analysis, for  
CC inhibiting protein function, or for tissue typing. The nucleic acid  
CC molecules are useful as probes, primers, chemical intermediates, or in  
CC biological assays. The present sequence represents a human enzyme  
CC protein.  
XX  
SQ Sequence 323 AA;

Query Match 100.0%; Score 1705; DB 5; Length 323;  
Best Local Similarity 100.0%; Pred. No. 7.5e-161;  
Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKSALFTRFFILLPWILIVIMLDVDTTRRPVPLTPRPYFSPYAVGGRGGLPLRRGGPA 60  
DB 1 MKSALFTRFFILLPWILIVIMLDVDTTRRPVPLTPRPYFSPYAVGGRGGLPLRRGGPA 60  
QY 61 HGTQKQNSRQPOPEPQLPTIYAITPTYSRPVQKAELTRLANTFRQVQLHWILVEDAA 120  
DB 61 HGTQKQNSRQPOPEPQLPTIYAITPTYSRPVQKAELTRLANTFRQVQLHWILVEDAA 120  
QY 121 ARSELVSRFLARAGLPSTHLHVPTPRYKRPGLPRATEQORNAGLWLRORHQRHQAQPGV 180  
DB 121 ARSELVSRFLARAGLPSTHLHVPTPRYKRPGLPRATEQORNAGLWLRORHQRHQAQPGV 180  
QY 181 LFFADDDNTYSLELFQEMRTTRKVSVPVGLVGGRRYERPLVGVGMYTGWADRPF 240  
DB 181 LFFADDDNTYSLELFQEMRTTRKVSVPVGLVGGRRYERPLVGVGMYTGWADRPF 240  
QY 241 AIDMAGFAVSLQVILSNPKAVFKRRGSGQPMQESDFLKQITTVVELEPKANNCTKVLVWH 300  
DB 241 AIDMAGFAVSLQVILSNPKAVFKRRGSGQPMQESDFLKQITTVVELEPKANNCTKVLVWH 300  
QY 301 TRTEKVNLANEPKYHLDTVKIEV 323  
DB 301 TRTEKVNLANEPKYHLDTVKIEV 323

RESULT 2  
ADQ97589  
ID ADQ97589 standard; protein; 323 AA.  
XX  
AC ADQ97589;  
XX  
DT 07-OCT-2004 (first entry)  
XX  
DE Human cancer associated sequence HP10-009, SEQ ID 566.  
XX  
KW Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human.  
XX  
OS Homo sapiens.  
XX  
FN WO2004060304-A2.  
XX  
PD 22-JUL-2004.  
XX  
PF 22-DEC-2003; 2003WO-US041389.  
XX  
PR 27-DEC-2002; 2002US-00330773.  
XX  
PA (SAGR-) SAGRES DISCOVERY INC.  
XX  
PI Morris DW, Malandro MS;  
XX  
DR WPI; 2004-543781/52.  
XX  
PT New isolated cancer associated nucleic acids comprising at least 10  
PT contiguous nucleotides, useful for diagnosing, preventing and/or treating  
PT cancers such as leukemia and lymphoma.  
XX

PS Claim 1; SEQ ID NO 566; 199pp; English.  
XX  
CC The present invention relates to cancer associated sequences (ADQ97025-  
CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or  
CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence  
CC data for this patent did not form part of the printed specification, but  
CC was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 323 AA;

Query Match 100.0%; Score 1705; DB 8; Length 323;  
Best Local Similarity 100.0%; Pred. No. 7.5e-161;  
Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKSALFTRFFILLPWILIVIMLDVDTTRRPVPLTPRPYFSPYAVGGRGGLPLRRGGPA 60  
DB 1 MKSALFTRFFILLPWILIVIMLDVDTTRRPVPLTPRPYFSPYAVGGRGGLPLRRGGPA 60  
QY 61 HGTQKQNSRQPOPEPQLPTIYAITPTYSRPVQKAELTRLANTFRQVQLHWILVEDAA 120  
DB 61 HGTQKQNSRQPOPEPQLPTIYAITPTYSRPVQKAELTRLANTFRQVQLHWILVEDAA 120  
QY 121 ARSELVSRFLARAGLPSTHLHVPTPRYKRPGLPRATEQORNAGLWLRORHQRHQAQPGV 180  
DB 121 ARSELVSRFLARAGLPSTHLHVPTPRYKRPGLPRATEQORNAGLWLRORHQRHQAQPGV 180  
QY 181 LFFADDDNTYSLELFQEMRTTRKVSVPVGLVGGRRYERPLVGVGMYTGWADRPF 240  
DB 181 LFFADDDNTYSLELFQEMRTTRKVSVPVGLVGGRRYERPLVGVGMYTGWADRPF 240  
QY 241 AIDMAGFAVSLQVILSNPKAVFKRRGSGQPMQESDFLKQITTVVELEPKANNCTKVLVWH 300  
DB 241 AIDMAGFAVSLQVILSNPKAVFKRRGSGQPMQESDFLKQITTVVELEPKANNCTKVLVWH 300  
QY 301 TRTEKVNLANEPKYHLDTVKIEV 323  
DB 301 TRTEKVNLANEPKYHLDTVKIEV 323

RESULT 3  
ADI03932  
ID ADI03932 standard; protein; 324 AA.  
XX  
AC ADI03932;  
XX  
DT 22-APR-2004 (first entry)  
XX  
DE Rat galactosidase beta-1,3-glucuronyltransferase.  
XX  
KW Enzyme protein; drug screening; disease diagnosis; rat; gene therapy;  
KW chromosome 6; enzyme; glucuronyltransferase;  
KW galactosidase beta-1,3-glucuronyltransferase.  
XX  
OS Rattus norvegicus.  
XX  
FN WO200268657-A2.  
XX  
PD 06-SEP-2002.  
XX  
PF 08-FEB-2002; 2002WO-US003623.  
XX  
PR 26-FEB-2001; 2001US-0270871P.  
PR 26-MAR-2001; 2001US-00816095.  
XX  
PA (PEXE) PE CORP NY.  
XX  
PI Gan W, Yan C, Merkulov GV, Di Francesco V, Beasley EM;  
XX  
DR WPI; 2002-713380/77.  
DR REFSEQ; NP\_072131.1.  
XX  
PT New human enzyme proteins, useful for treating or diagnosing disorders

PT associated with abnormal expression of the protein, in drug screening  
PT assays and pharmacogenomic analysis.

PS Disclosure; SEQ ID NO 4; 127pp; English.

XX The invention relates to a novel isolated enzyme protein and encoding  
XX polynucleotides. The protein shows a high degree of similarity to a  
XX glucuronyltransferase cloned from a rabbit brain cDNA library. The  
XX peptides and nucleic acid molecules are useful as models for the  
XX development of human therapeutic targets, aid in the identification of  
XX therapeutic proteins, and serve as targets for the development of human  
XX therapeutic agents. The peptide may be used in drug screening assays, in  
XX assays to determine the biological activity of the protein, to raise  
XX antibodies or to elicit another immune response, as a reagent in assays  
XX designed to quantitatively determine levels of the protein in biological  
XX fluids, or as markers for tissues in which the corresponding protein is  
XX preferentially expressed. The human enzyme protein is also useful for  
XX diagnosing a disease, predisposition to a disease, or treating a disorder  
XX characterized by an absence of, inappropriate or unwanted expression of  
XX the protein. The antibodies are useful in pharmacogenomic analysis, for  
XX inhibiting protein function, or for tissue typing. The nucleic acid  
XX molecules are useful as probes, primers, chemical intermediates, or in  
XX biological assays. The present sequence represents a rat galactosidase  
XX beta-1,3-glucuronyltransferase used in alignment studies with the novel  
XX human enzyme protein.

SQ Sequence 324 AA;

Query Match 89.5%; Score 1526.5; DB 5; Length 324;  
Best Local Similarity 89.8%; Pred. No. 4.6e-143;  
Matches 291; Conservative 10; Mismatches 22; Indels 1; Gaps 1;

QY 1 M K S A L T R F F I L L P W I L I V I I M L D V T R P V P L T R P Y F S P V A V G G G A R L P L R R G G P A 60  
DB 1 M K S A L C N R F F I L L P W I L I V I I M L D V D P R P A P Q L T R P Y F S P V T V G C G S R V P L R R S S P G 60  
QY 61 H - G T Q K R N S R P Q P E P O L P T T I Y A I T P T Y S R P V Q K A E L T R L A N T F R Q V A Q L H W I L V E D A 119  
DB 61 R D A E K R N S R P Q L P E P L T I Y A I T P T Y S R P V Q K A E L T R L A N T F R Q V A Q L H W I L V E D R 120  
QY 120 A A R S E L V S R F L A R A G L P S T H L V P T R R Y K R P L P R A T E Q R N A G L A W L R Q R H Q H Q A Q P G 179  
DB 121 A T R S E L V S S F L A R A G L P N T H L V P T R R Y K R P L P R A T E Q R N A G L A W L R Q R H Q H Q A Q P G 180  
QY 180 V L F F A D D D N T Y S L E L F Q E M T T R K V S W P V G L V G G R R Y E R P L V E N G K V G W Y T G W A D R P 239  
DB 181 V L F F A D D D N T Y S L E L F Q E M T T R K V S W P V G L V G G R R Y E R P L V K N G K V G W Y T G W E D R P 240  
QY 240 F A I D M A G F A V S L Q V I L S N P K A V F K R G S Q P G M O E S D F L K Q I T T V E E L E P K A N N C T K V L W 299  
DB 241 F A I D M A G F A V S L Q V I L S N P K A V F K R G S Q P G M O E S D F L K Q I T T V E E L E P K A N N C T K V L W 300  
QY 300 H T R T E K V N L A N E P K Y H L D T V K I E V 323  
DB 301 H T R T E K V N L A N E P K Y H M D T V N I E V 324

RESULT 4  
ADQ97586  
ID ADQ97586 standard; protein; 292 AA.

XX AC ADQ97586;  
XX DT 07-OCT-2004 (first entry)  
XX DE Mouse cancer associated sequence MP10-009, SEQ ID 563.  
XX KW Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse.  
XX OS Mus musculus.  
XX PN WO2004060304-A2.  
XX XX

PD 22-JUL-2004.  
XX 22-DEC-2003; 2003WO-US041389.

XX 27-DEC-2002; 2002US-00330773.

XX (SAGR-) SAGRES DISCOVERY INC.

XX Morris DW, Malandro MS;

XX WPI; 2004-543781/52.

XX New isolated cancer associated nucleic acids comprising at least 10  
XX contiguous nucleotides, useful for diagnosing, preventing and/or treating  
XX cancers such as leukemia and lymphoma.

XX Claim 1; SEQ ID NO 563; 199pp; English.

XX The present invention relates to cancer associated sequences (ADQ97025-  
XX ADQ98004). The sequences are useful for the diagnosis, prevention and/or  
XX treatment of cancer, such as leukemia and lymphoma. Note: The sequence  
XX data for this patent did not form part of the printed specification, but  
XX was obtained in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 292 AA;

Query Match 71.7%; Score 1223; DB 8; Length 292;  
Best Local Similarity 74.1%; Pred. No. 6.9e-113;  
Matches 240; Conservative 9; Mismatches 17; Indels 58; Gaps 2;

QY 1 M K S A L T R F F I L L P W I L I V I I M L D V T R P V P L T R P Y F S P V A V G G G A R L P L R R G G P A 60  
DB 26 M K S A L C S R F F I L L P W I L I V I I M L D V D P R P A P Q L T R P Y F S P H A V G C G S R V P L R S S P G 85  
QY 61 H - G T Q K R N S R P Q P E P O L P T T I Y A I T P T Y S R P V Q K A E L T R L A N T F R Q V A Q L H W I L V E D A 119  
DB 86 R D A E K R N S R P Q L P E P L T I Y A I T P T Y S R P - - - - - 118  
QY 120 A A R S E L V S F L A R A G L P S T H L V P T R R Y K R P L P R A T E Q R N A G L A W L R Q R H Q H Q A Q P G 179  
DB 119 - - - - - R P W L P R A T E Q R N A G L A W L R Q R H Q H Q A Q P G 148  
QY 180 V L F F A D D D N T Y S L E L F Q E M T T R K V S W P V G L V G G R R Y E R P L V E N G K V G W Y T G W A D R P 239  
DB 149 V L F F A D D D N T Y S L E L F Q E M T T R K V S W P V G L V G G R R Y E R P L V K N G K V G W Y T G W E D R P 208  
QY 240 F A I D M A G F A V S L Q V I L S N P K A V F K R G S Q P G M O E S D F L K Q I T T V E E L E P K A N N C T K V L W 299  
DB 209 F A I D M A G F A V S L Q V I L S N P K A V F K R G S Q P G M O E S D F L K Q I T T V E E L E P K A N N C T K V L W 268  
QY 300 H T R T E K V N L A N E P K Y H L D T V K I E V 323  
DB 269 H T R T E K V N L A N E P K Y H L D T V N I E V 292

RESULT 5  
AD103933  
ID AD103933 standard; protein; 197 AA.

XX AC AD103933;  
XX DT 22-APR-2004 (first entry)  
XX DE Human UDP-glucuronyltransferase-S fragment.  
XX KW Enzyme protein; drug screening; disease diagnosis; rat; gene therapy;  
XX chromosome 6; enzyme; glucuronyltransferase; UDP-glucuronyltransferase-S.  
XX OS Homo sapiens.  
XX PN WO200268657-A2.  
XX XX

PD 06-SEP-2002.  
XX  
PF 08-FEB-2002; 2002WO-US003623.  
XX  
PR 26-FEB-2001; 2001US-0270871P.  
PR 26-MAR-2001; 2001US-00816095.  
XX  
XX (PEKE ) PE CORP NY.  
XX  
XX Gan W, Yan C, Merkulov GV, Di Francesco V, Beasley EM;  
XX WPI; 2002-713380/77.  
XX  
XX New human enzyme proteins, useful for treating or diagnosing disorders  
PT associated with abnormal expression of the protein, in drug screening  
PT assays and pharmacogenomic analysis.  
XX  
XX Disclosure; SEQ ID NO 5; 127pp; English.  
XX  
XX The invention relates to a novel isolated enzyme protein and encoding  
CC polynucleotides. The protein shows a high degree of similarity to a  
CC glucuronyltransferase cloned from a rabbit brain cDNA library. The  
CC peptides and nucleic acid molecules are useful as models for the  
CC development of human therapeutic targets, aid in the identification of  
CC therapeutic proteins, and serve as targets for the development of human  
CC therapeutic agents. The peptide may be used in drug screening assays, in  
CC assays to determine the biological activity of the protein, to raise  
CC antibodies or to elicit another immune response, as a reagent in assays  
CC designed to quantitatively determine levels of the protein in biological  
CC fluids, or as markers for tissues in which the corresponding protein is  
CC preferentially expressed. The human enzyme protein is also useful for  
CC diagnosing a disease, predisposition to a disease, or treating a disorder  
CC characterized by an absence of, inappropriate or unwanted expression of  
CC the protein. The antibodies are useful in pharmacogenomic analysis, for  
CC inhibiting protein function, or for tissue typing. The nucleic acid  
CC molecules are useful as probes, primers, chemical intermediates, or in  
CC biological assays. The present sequence represents a human UDP-  
CC glucuronyltransferase-S fragment used in alignment studies with the novel  
XX human enzyme protein.  
XX  
SQ Sequence 197 AA;

Query Match 60.9%; Score 1039; DB 5; Length 197;  
Best Local Similarity 100.0%; Pred. No. 8.5e-95;  
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKSALTRFRFILLPWILLIIMLDVTRRPVPLTPRPYFSPYAVGGRGARLPRLRGGA 60  
DB 1 MKSALTRFRFILLPWILLIIMLDVTRRPVPLTPRPYFSPYAVGGRGARLPRLRGGA 60  
QY 61 HGTKRNQSRPQPEPQPLPTIYAITPTYSRPVQKAELTRLANTPRQVQLHWILVEDAA 120  
DB 61 HGTKRNQSRPQPEPQPLPTIYAITPTYSRPVQKAELTRLANTPRQVQLHWILVEDAA 120  
QY 121 ARSELVSRFLARAGLSTHLHVTTPRYKRPGLPRATEQGNAGLAWLRQHQHQAQPGV 180  
DB 121 ARSELVSRFLARAGLSTHLHVTTPRYKRPGLPRATEQGNAGLAWLRQHQHQAQPGV 180  
QY 181 LFFADDNDNTYSELFOE 197  
DB 181 LFFADDNDNTYSELFOE 197

RESULT 6  
ABB97603  
ID ABB97603 standard; protein; 347 AA.  
XX  
AC ABB97603;  
XX  
XX 27-JUN-2002 (first entry)  
XX  
XX Novel human protein SEQ ID NO: 871.

KW Human; antianaemic; vulnery; antiinflammatory; immunomodulator;  
KW antifertility; cerebroprotective; cytostatic; rheumatic; gene therapy;  
KW neuroprotective; antiparkinsonian; protein therapy; EST;  
XX expressed sequence tag.  
OS Homo sapiens.  
XX WO200222660-A2.  
XX  
XX 21-MAR-2002.  
XX  
XX 10-SEP-2001; 2001WO-US026015.  
XX  
XX 11-SEP-2000; 2000US-00659671.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
XX Xue AJ, Yang Y, Wehrman T, Drmanac RT;  
XX  
XX WPI; 2002-292408/33.  
XX N-PSDB; ABN32789.  
XX  
XX An isolated polynucleotide for treating diseases associated with its  
PT encoded polypeptide such as cancer and multiple sclerosis.  
XX  
XX Example 2; SEQ ID NO 871; 509pp; English.  
XX  
XX The present invention provides the protein and coding sequences of 444  
CC novel human proteins. These were isolated from expressed sequences tags  
CC (ESTs). They can be used to stimulate cell growth, to regulate  
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth  
CC e.g. in burn treatment, to regulate the immune system e.g. to treat  
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat  
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke  
CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.  
CC rheumatoid arthritis, and to treat nervous system disorders e.g.  
CC Parkinson's disease. The present sequence is a protein of the invention  
XX  
SQ Sequence 347 AA;

Query Match 46.2%; Score 787.5; DB 5; Length 347;  
Best Local Similarity 49.0%; Pred. No. 2.2e-69;  
Matches 169; Conservative 37; Mismatches 84; Indels 55; Gaps 8;  
QY 11 ILLPWILIVII-----MLDV-----DTRRPVPP-LTPRPYFSPYAVGGRGARLPRLR 55  
DB 26 IVLPWTLITVWQSTLAPLAVHKDEGSDPRRTPPGADPREYC----- 70  
QY 56 RGGPAHGTQKRN-----QSRPQPEPQPLPTIYAITPTYSRPVQKAELTRLANTF 105  
DB 71 -----TSDRIVEVWRTEYVYTRPPWSD-TLPTIHVVVTPYSRPVQKAELTRMANTL 122  
QY 106 ROVAOLHWILVEDAAARSELVSRFLARAGLSTHLHVTTPRYK-----RPLPRATEQ 159  
DB 123 LHVPLNLHWLWEDAPRRTPPLTARLLRDTGLNTHLVETPRNYKLRGDARDPRIPRGTNQ 182  
QY 160 RNAGLAWLRQHQHQAQPGVLFADDNDNTYSELFOEMRTTRKYSVWPVGLVGRRYER 219  
DB 183 RNALRWLRETTPRNSQPGVYVYFADDNDNTYSELFOEMRSTRYSVWPVAVGGLRYEA 242  
QY 220 PLVEN-GKVGWYGTWRADRPFAIDMAGFAVLSQVILNPNKAVFKRRSQPGKQSDFLK 278  
DB 243 PRVNGAGKVGWKTVDPRHPRFAIDMAGFAVNLRLILQSRQAYFKLRGKGYQESSLLR 302  
QY 279 QITTVEELEPKANCTKVLVWHTRTKYNLANEPKYVLDLTVKIEV 323  
DB 303 ELVTINDLEPKAANCTKILVWHTRTKEPKVLENEKKGGTDPDSVEI 347  
RESULT 7  
ABO07264  
ID ABO07264 standard; protein; 334 AA.



SQ	Sequence 347 AA;	
Query Match	45.6%; Score 777.5; DB 2; Length 347;	
Best Local Similarity	48.3%; Pred. No. 2.2e-68;	
Matches 167; Conservative	36; Mismatches 86; Indels 57; Gaps 8;	
QY	11 ILLPWILIVII-----MLDV-----DTRRPVPP-LTPRPYF-----SPYA 44	
DB	26 IVLPTLLITVWHQSSLAFLAVHKDEGSDPRHEAPPGADPREYCMDSRDDIVEVVRTEYV 85	
QY	45 VGRGGARLPLRRGGPAHGTCQKNSRQPOPEPOLPTIYAITPTYSRPVQKAELTRANT 104	
DB	86 -----YTRPPWSD-TLPTIHVVTTPTYSRPVQKAELTRANT 121	
QY	105 FRQVAQLHWILVEDAAARSELVSRLARAGLSPSTHLHVPTRRYK-----RPGLPRATE 158	
DB	122 LLHVPNLHWLVVEDAPRRTPLTARLLRDTGLNYTHLVETPRNYKLRGDARDPRI PRGTM 181	
QY	159 QRNAGLAWLRQHQHQAQPGVLFADDDNTYSLELFQEMRTTRKVSVPVGLVGGRRYE 218	
DB	182 QRLALRWLRETFFRNSSTQPGVYFADDDNTYSLELFEEMRSTRVSVWVPAFVGLGYE 241	
QY	219 RPLVEN-GKVVGWYTGWRADRPFAIDMAGFAVSLQVILSNPKAVFKRGSGQPCWQSSDFL 277	
DB	242 APRVNGAGKVVGWKTVDPHRPFADIDMAGFAVNLRLILQSRQAYFKLRGVKGYQESSLL 301	
QY	278 KOITVVELEPKANNCTKVLVWHTREKVNLANEPKYHLDTVKIEV 323	
DB	302 RELVTLNDLEPKAANCTKILVWHTREKPVLVNEGKKGFTDPSVEI 347	
RESULT 9		
AAW90057	AAW90057 standard; protein; 347 AA.	
XX	AC AAW90057;	
XX	DT 26-FEB-1999 (first entry)	
XX	DE Rat GlcAT-P protein.	
XX	KW GlcAT-P; glucuronate transferase; rat; glucuronic acid; donor; receptor.	
XX	OS Rattus sp.	
XX	PN JP10313867-A.	
XX	PD 02-DEC-1998.	
XX	PF 16-MAY-1997; 97JP-00127065.	
XX	PR 16-MAY-1997; 97JP-00127065.	
XX	PA (SEKG ) SEIKAGAKU KOGYO CO LTD.	
XX	WPI; 1999-074149/07.	
XX	N-PSDB; AAW73869.	
XX	DNA coding glucuronate transferase - comprises transferring glucuronic acid from glucuronic acid donor to glucuronic acid receptor.	
XX	Claim 3; Page 13-14; 18pp; Japanese.	
XX	This sequence represents a novel rat glucuronate transferase (GlcAT-P) which transfers glucuronic acid from a glucuronic acid donor to a glucuronic acid receptor with an optimum reaction pH of 6.0 to 6.5. The protein transfers specifically glucuronic acid to the N-acetylglucosamine residue of asialo-orosomucoid and neuron adhesive molecule, has a molecular weight of approximately 45,000 and maintains its activity in the presence of 5 mM neolactotetraose-phenylC14H29	
XX		
SQ	Sequence 347 AA;	
Query Match	45.6%; Score 777.5; DB 2; Length 347;	
Best Local Similarity	48.3%; Pred. No. 2.2e-68;	
Matches 167; Conservative	36; Mismatches 86; Indels 57; Gaps 8;	
QY	11 ILLPWILIVII-----MLDV-----DTRRPVPP-LTPRPYF-----SPYA 44	
DB	26 IVLPTLLITVWHQSSLAFLAVHKDEGSDPRHEAPPGADPREYCMDSRDDIVEVVRTEYV 85	
QY	45 VGRGGARLPLRRGGPAHGTCQKNSRQPOPEPOLPTIYAITPTYSRPVQKAELTRANT 104	
DB	86 -----YTRPPWSD-TLPTIHVVTTPTYSRPVQKAELTRANT 121	
QY	105 FRQVAQLHWILVEDAAARSELVSRLARAGLSPSTHLHVPTRRYK-----RPGLPRATE 158	
DB	122 LLHVPNLHWLVVEDAPRRTPLTARLLRDTGLNYTHLVETPRNYKLRGDARDPRI PRGTM 181	
QY	159 QRNAGLAWLRQHQHQAQPGVLFADDDNTYSLELFQEMRTTRKVSVPVGLVGGRRYE 218	
DB	182 QRLALRWLRETFFRNSSTQPGVYFADDDNTYSLELFEEMRSTRVSVWVPAFVGLGYE 241	
QY	219 RPLVEN-GKVVGWYTGWRADRPFAIDMAGFAVSLQVILSNPKAVFKRGSGQPCWQSSDFL 277	
DB	242 APRVNGAGKVVGWKTVDPHRPFADIDMAGFAVNLRLILQSRQAYFKLRGVKGYQESSLL 301	
QY	278 KOITVVELEPKANNCTKVLVWHTREKVNLANEPKYHLDTVKIEV 323	
DB	302 RELVTLNDLEPKAANCTKILVWHTREKPVLVNEGKKGFTDPSVEI 347	
RESULT 10		
ABO07265	ABO07265 standard; protein; 335 AA.	
XX	AC ABO07265;	
XX	DT 13-AUG-2003 (first entry)	
XX	DE Human p53 modifying protein, SEQ ID 225.	
XX	KW Human; p53 modifier; cytostatic; cancer; cytostatic; antiangiogenic; antipapillary; p53 pathway; breast cancer; colon cancer; kidney cancer; lung cancer; ovarian cancer; angiogenesis; cell cycle; apoptotic disorder; cell proliferation disorder.	
XX	OS Homo sapiens.	
XX	PN WO200299122-A1.	
XX	PD 12-DEC-2002.	
XX	PF 03-JUN-2002; 2002WO-US017382.	
XX	PR 05-JUN-2001; 2001US-0296076P.	
XX	PR 10-OCT-2001; 2001US-0328605P.	
XX	PR 15-FEB-2002; 2002US-0357253P.	
XX	(EXEL-) EXELIXIS INC.	
XX	Friedman L, Plozman GD, Belvin M, Francis-Lang H, Li D, Funke RP;	
XX	WPI; 2003-156859/15.	
XX	N-PSDB; ACD13438.	
XX	Identifying modulators of the p53 pathway for use in treating apoptotic or cell proliferation disorders, comprises screening for agents that modulate activity of a human ortholog of genes that modify the p53 pathway in Drosophila.	
XX	Example 2; Page 662-663; 678pp; English.	
XX	The invention relates to identifying (M1) a candidate p53 pathway	



XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.  
KW Homo sapiens.  
XX WO2004048938-A2.  
XX 10-JUN-2004.  
XX 26-NOV-2003; 2003WO-US038193.  
XX 26-NOV-2002; 2002US-0429739P.  
XX (PROT-) PROTEIN DESIGN LABS INC.  
XX Aziz N, Ginsburg WM, Zlotnik A;  
XX WPI; 2004-441208/41.  
XX Early detection of soft tissue sarcoma comprises determining expression  
PT of a gene in a first soft tissue sample and a normal soft tissue sample  
PT and comparing the gene expression, also useful in treating soft tissue  
PT sarcoma.  
XX  
XX Example 2; SEQ ID NO 4238; 210pp; English.

XX The invention relates to a novel method for detecting soft tissue sarcoma  
CC which comprises obtaining a first soft tissue sample from an individual  
CC and a normal soft tissue sample from the same or different individual,  
CC determining the expression of a gene in both samples and comparing the  
CC expression of the gene in both soft tissue samples, where a higher level  
CC of protein expression in the first soft tissue sample indicates the  
CC presence of soft tissue sarcoma. The method of the invention has  
CC cytotatic applications and may be useful for detecting soft tissue  
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic  
CC acid sequences may be useful in diagnostic and screening applications.  
CC The current sequence is that of a human soft tissue sarcoma-upregulated  
CC protein of the invention. The current sequence is not shown within the  
CC specification per se but was submitted in CD format by the inventor.

XX Sequence 335 AA;

Query Match 38.3%; Score 653; DB 8; Length 335;  
Best Local Similarity 43.9%; Pred. No. 5.3e-56;  
Matches 150; Conservative 41; Mismatches 93; Indels 58; Gaps 10;

QY 1 MKSALFTRFFILLPWILVIMLDVTRRP---VPPLTPRPYSPYAVGRGGA----- 50  
DB 5 LKNVFLAYFLVSIAGLLYALVOLG---QPCDCLPPL-----RAAAEQLRQKD 48  
QY 51 -----RLPLRRGGPAHGTQKRNQSRPQPQPEQ-LPTIYAITPTYSRPVQKAELTRLAN 103  
DB 49 LRISQLQAEULRRPPA-----PAQPEPEALPTIVVTTVARLVQKAEVLRSQ 98  
QY 104 TFRQVLAQLHWILVEDAAASSELVSRFLARAGLSTHLVTPR----RYKRPGL--PRAT 157  
DB 99 TSLVPLRLHLLVEDAEGPTPLVSGLLAASGLLFTLVLTPRAQLREGEQVHWPRGV 158  
QY 158 EQRNAGLAWLR-----ORHQHQAQPGVLFADDDNTYSLELFQEMRTTRKVSVPV 209  
DB 159 EQRNKALDWLRGGVAGGKDPPTGQTQGVTFADDDNTYSLELFQEMRWTRKVSVPV 218  
QY 210 GLVGGRRYERPLVGVGWYTGWRADRPFAIDMAGFVLSQVILSNPKAVFKRRGSQP 269  
DB 219 GLVGLGRFEGPQVQDGRVGVGHTAWEPSPRPFPVDMAGFVALPLLDKPNQAFDSTAPR- 277  
QY 270 GMESEDFLKQITVVELEPKANNCTKVLVWHTTRTEKVNLANE 311  
DB 278 GHLESSLLSLVDPKOLEPRAANCTRVLVWHTTRTEKPKMQE 319

RESULT 13  
ADY70386

ID ADY70386 standard; protein; 335 AA.  
XX ADY70386;  
AC  
XX  
DT 02-JUN-2005 (first entry)  
XX  
DE Galactosylgalactosylxylosylprotein 3-beta-glucuronosyltransferase 3.  
XX  
KW beta-amyloid; amyloid precursor protein; neuroprotective; nootropic;  
KW neurodegenerative disease; Alzheimers disease.  
XX  
OS Homo sapiens.  
XX  
PN WO2005023858-A1.  
XX  
PD 17-MAR-2005.  
XX  
PF 10-DEC-2003; 2003WO-EP013980.  
XX  
PR 05-SEP-2003; 2003EP-00019642.  
XX  
XX (CELL-) CELLZONE AG.  
PA  
XX Bouwmeester T, Drewes G, Hopf C, Joberty G, Rowley A;  
PI  
XX WPI; 2005-242127/25.  
DR  
XX  
XX Protein complex of beta-amyloid precursor protein processing pathway.  
PT useful as target for active agent of pharmaceutical, in treating or  
PT preventing Alzheimer's disease and related neurodegenerative disorders.  
XX  
PS Claim 1; SEQ ID NO 118; 1294pp; English.  
XX  
XX The invention relates to novel protein complexes of the beta-amyloid  
CC precursor protein processing pathway. The invention further comprises:  
CC component proteins of said complexes, fragments and derivatives of the  
CC component proteins and antibodies specific to the complexes. The  
CC invention also relates to methods for the use of the complexes of the APP  
CC processing pathway and their interacting proteins in screening, diagnosis  
CC and therapy, as well as to methods for preparing the complexes. The novel  
CC protein complexes have neuroprotective and nootropic activities. The  
CC protein complex, protein chosen from the proteins of Table 1 and  
CC antibodies are useful for diagnosing diseases and disorders,  
CC preferentially for diseases or disorders such as neurodegenerative  
CC disease such as Alzheimer's disease and related neurodegenerative  
CC disorders. The method is useful for treating or preventing a disease or  
CC disorder characterized by an aberrant amount of, activity of, component  
CC composition of or intracellular localization of the protein complex. The  
CC disease or disorder involves decreased levels of the amount or activity  
CC of the complex, or increased levels of the amount or activity of the  
CC protein complex. The protein complex and/or proteins listed in table 1,  
CC fifth column of the complex, is useful as a target for an active agent of  
CC a pharmaceutical, preferably a drug target in the treatment or prevention  
CC of a disease or disorder, preferentially of a disease or disorder such as  
CC neurodegenerative disease such as Alzheimer's disease and related  
CC neurodegenerative disorders. The pharmaceutical composition and method  
CC are useful for treating diseases and disorders, preferentially for  
CC diseases or disorders such as neurodegenerative disease such as  
CC Alzheimer's disease and related neurodegenerative disorders. This  
CC sequence represents one of the proteins used in a complex of the beta-  
CC amyloid precursor protein processing pathway of the invention.

XX Sequence 335 AA;

Query Match 37.8%; Score 645; DB 9; Length 335;  
Best Local Similarity 43.6%; Pred. No. 3.3e-55;  
Matches 149; Conservative 41; Mismatches 94; Indels 58; Gaps 10;

QY 1 MKSALFTRFFILLPWILVIMLDVTRRP---VPPLTPRPYSPYAVGRGGA----- 50  
DB 5 LKNVFLAYFLVSIAGLLYALVOLG---QPCDCLPPL-----RAAAEQLRQKD 48  
QY 51 -----RLPLRRGGPAHGTQKRNQSRPQPQPEQ-LPTIYAITPTYSRPVQKAELTRLAN 103



Qy	77	PQLPTIYAITPTYSRPVQKAELTRLANTFRQVAQLHWILVEDAAARSELVSRFLARAGLP	136
Db	5	PPMPGMY-----KAEVLRLSQTLSLVRPLHLLVEDAEGPTPLVSGLLAASGLL	53
Qy	137	STHLHVTPR----RYKRPGI--PRATEORNAGLAWLR-----QRHQHQAQPGVLF	182
Db	54	FTHLVVLTPKAQRLREGEPCGWHPRGVEQRNKALDWLRGRGGAVGGEKDPPTPGTQGVVY	113
Qy	183	FADDDNTYSLELFEQEMRTTRKVSVPVGLVGGRRYERPLVENGKVVGYWTGWRADRPFAI	242
Db	114	FADDDNTYSRELFEEMRWTRGVSVMPVGLVGLRFEFGPQVQDGRVVGFHTAWEPSRPFV	173
Qy	243	DMAGFAVSLQVILSNPKAVFKRRGSPQMGQESDFLKQITTVBELEPKANNCTKVLVWHTR	302
Db	174	DMAGFAVALPPLLKPKNAQFDSTAPR-GHLESLLSHLVDPKDLEPRAANCTRSLAVSPR	232
Qy	303	TE	304
Db	233	LE	234

Search completed: May 9, 2006, 10:53:26  
Job time : 90 secs

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2006, 10:54:58 ; Search time 47 Seconds  
(without alignments)  
568.175 Million cell updates/sec

Title: US-10-634-905-2  
Perfect score: 1705  
Sequence: 1 MKSALTRFFILLPWILVI.....EKNVLANEPKYHLDTVKIEV 323

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Issued Patents AA.\*
- 1: /cgn2\_6/ptodata/1/iaa/5\_COMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/iaa/6\_COMB.pep.\*
  - 3: /cgn2\_6/ptodata/1/iaa/H\_COMB.pep.\*
  - 4: /cgn2\_6/ptodata/1/iaa/PTUS\_COMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/iaa/RE\_COMB.pep.\*
  - 6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1705	100.0	323	2	US-09-816-095-2
2	1526.5	89.5	324	2	US-09-816-095-4
3	1039	60.9	197	2	US-09-816-095-5
4	777.5	45.6	347	2	US-09-059-369-2
5	486.5	28.5	291	2	US-09-270-767-33192
6	486.5	28.5	291	2	US-09-270-767-48409
7	236	13.8	142	2	US-09-270-767-45263
8	130.5	7.7	521	2	US-09-252-991A-18119
9	115.5	6.8	603	2	US-09-252-991A-25264
10	111	6.5	666	2	US-09-050-739-70
11	107	6.3	396	2	US-09-252-991A-26907
12	107	6.3	1037	2	US-09-252-991A-25361
13	106.5	6.2	475	2	US-09-252-991A-25837
14	106.5	6.2	720	2	US-09-252-991A-23212
15	105.5	6.2	388	2	US-09-252-991A-17917
16	104	6.1	470	2	US-09-252-991A-19467
17	104	6.1	618	2	US-09-252-991A-23696
18	103	6.0	25	2	US-09-059-369-4
19	102	6.0	1418	2	US-09-252-991A-33367
20	99.5	5.8	436	2	US-09-252-991A-20220
21	99.5	5.8	506	2	US-09-252-991A-18165
22	99.5	5.8	560	2	US-09-252-991A-26107
23	99	5.8	361	2	US-09-252-991A-24065
24	98	5.7	288	2	US-09-252-991A-27950
25	97.5	5.7	1706	2	US-09-252-991A-31760
26	97	5.7	485	2	US-09-252-991A-17141
27	97	5.7	618	2	US-09-252-991A-28358

ALIGNMENTS

RESULT 1

US-09-816-095-2  
; Sequence 2, Application US/09816095  
; Patent No. 6664084  
; GENERAL INFORMATION:  
; APPLICANT: GAN, Weiniu  
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES  
; FILE REFERENCE: CLO01147  
; CURRENT APPLICATION NUMBER: US/09/816,095  
; CURRENT FILING DATE: 2001-03-26  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 323  
; TYPE: PRT  
; ORGANISM: Human  
US-09-816-095-2

Query Match	100.0%;	Score 1705;	DB 2;	Length 323;
Best Local Similarity	100.0%;	Pred. No. 7e-174;		
Matches 323;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MKSALTRFFILLPWILVIIMLDVTRRPVPLTPRPYFSPYAVGCGARLPLRRGGA	60	Sequence 23317, A
DB	1	MKSALTRFFILLPWILVIIMLDVTRRPVPLTPRPYFSPYAVGCGARLPLRRGGA	60	Sequence 24741, A
QY	61	HGTOKENQSRPOPEPOLPTIYAITPTYSRPVQKAELTRLANTFRVQALHWILVEDAA	120	Sequence 26509, A
DB	61	HGTOKENQSRPOPEPOLPTIYAITPTYSRPVQKAELTRLANTFRVQALHWILVEDAA	120	Sequence 22313, A
QY	121	ARSELVSRFLAAGLPSTHLVPTPRYKRPGLPRATEQORNAGLWLRQHQHQRQPGV	180	Sequence 20216, A
DB	121	ARSELVSRFLAAGLPSTHLVPTPRYKRPGLPRATEQORNAGLWLRQHQHQRQPGV	180	Sequence 24173, A
QY	181	LPFADDNTYSLELFOEMRTTRKVSVPVGLVGGRRYERPLVENGKVGVYTGWRADRP	240	Sequence 20345, A
DB	181	LPFADDNTYSLELFOEMRTTRKVSVPVGLVGGRRYERPLVENGKVGVYTGWRADRP	240	Sequence 19431, A
QY	241	AIDMAGFVSLQVILSNPKAVFKRGSPQMDESDFLKQITTVVELEPKANNCTKVLVWH	300	Sequence 18960, A
DB	241	AIDMAGFVSLQVILSNPKAVFKRGSPQMDESDFLKQITTVVELEPKANNCTKVLVWH	300	Sequence 29050, A
QY	301	TRTEKVNLANEPKYHLDTVKIEV 323		Sequence 28585, A
DB	301	TRTEKVNLANEPKYHLDTVKIEV 323		Sequence 25366, A

RESULT 2

```

US-09-816-095-4
; Sequence 4, Application US/09816095
; Patent No. 6664084
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001147
; CURRENT APPLICATION NUMBER: US/09/816, 095
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-816-095-4

Query Match      89.5%; Score 1526.5; DB 2; Length 324;
Best Local Similarity 89.8%; Pred. No. 8.9e-155;
Matches 291; Conservative 10; Mismatches 22; Indels 1; Gaps 1;

QY 1 MKSALTRFFILLPWILVIMLDVDTTRRPVPLTPRPYFSPYAVGRGARLPRLRGGA 60
DB 1 MKSALCNRFILLPWILVIMLDVDTTRRPVPLTPRPYFSPYAVGRGARLPRLRGGA 60
QY 61 H-GTQKRNQSRPQPEPQLPTIYAITPTYSRVPQKAELTRANTFRQVAQLHWILVEDAA 119
DB 61 RDAAEKNESRPLQPEPRLPTIYAITPTYSRVPQKAELTRANTFRQVAQLHWILVEDR 120
QY 120 AARSELVSRFLARAGLPSTHLHVPTPRRYKRPGLPRATEQORNAGLAWLRQHQRHQAQPG 179
DB 121 ATRSELVSSFLARAGLPSTHLHVPTPRRYKRPGLPRATEQORNAGLAWLRQHQRHQAQPG 180
QY 180 VLFADDNTYSLELFOEHTTKVSWPVLGGRYRPLVNGKVGWYTGWRADRP 239
DB 181 VLFADDNTYSLELFOEHTTKVSWPVLGGRYRPLVNGKVGWYTGWRADRP 240
QY 240 FAIDMAGFAVSLQVILSNKAVFKRGSGQGMQESDFLKQITTVDELEPKANNCTKVLVW 299
DB 241 FAIDMAGFAVSLQVILSNKAVFKRGSGQGMQESDFLKQITTVDELEPKANNCTKVLVW 300
QY 300 HTRTEKVNLANEPKYHLDTVKIEV 323
DB 301 HTRTEKVNLANEPKYHMDTVNIEV 324

RESULT 3
US-09-816-095-5
; Sequence 5, Application US/09816095
; Patent No. 6664084
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001147
; CURRENT APPLICATION NUMBER: US/09/816, 095
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Human
US-09-816-095-5

Query Match      60.9%; Score 1039; DB 2; Length 197;
Best Local Similarity 100.0%; Pred. No. 6.3e-103;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKSALTRFFILLPWILVIMLDVDTTRRPVPLTPRPYFSPYAVGRGARLPRLRGGA 60
DB 1 MKSALTRFFILLPWILVIMLDVDTTRRPVPLTPRPYFSPYAVGRGARLPRLRGGA 60
QY 61 HGTQKRNQSRPQPEPQLPTIYAITPTYSRVPQKAELTRANTFRQVAQLHWILVEDAA 120
DB 61 HGTQKRNQSRPQPEPQLPTIYAITPTYSRVPQKAELTRANTFRQVAQLHWILVEDAA 120
QY 121 ARSELVSRFLARAGLPSTHLHVPTPRRYKRPGLPRATEQORNAGLAWLRQHQRHQAQPGV 180
DB 121 ARSELVSRFLARAGLPSTHLHVPTPRRYKRPGLPRATEQORNAGLAWLRQHQRHQAQPGV 180
QY 181 LFFADDNTYSLELFOE 197
DB 181 LFFADDNTYSLELFOE 197

US-09-059-369-2
; Sequence 2, Application US/09059369
; Patent No. 6040156
; GENERAL INFORMATION:
; APPLICANT: KAWASAKI, TOSHISUKE
; APPLICANT: OKA, SHOGO
; TITLE OF INVENTION: DNA ENCODING GLUCURONYLTRANSFERASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSES: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/059,369
; FILING DATE: 14-APR-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 9-127065
; FILING DATE: 16-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 9378-0002-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-059-369-2

Query Match      45.6%; Score 777.5; DB 2; Length 347;
Best Local Similarity 48.3%; Pred. No. 1.4e-74;
Matches 167; Conservative 36; Mismatches 86; Indels 57; Gaps 8;

QY 11 ILLPWILVII-----MLDV-----DTRRPVPP-LTRPYF-----SPYA 44
DB 26 IVLPWTILITVWHQSSLAPLLAVHDKGSDPRHEAPPGADPREYCMRDIDVEVVRTEV 85
QY 45 VGRGGARLPRLRGGAHGTQKRNQSRPQPEPQLPTIYAITPTYSRVPQKAELTRANT 104
DB 86 -----YTRPPWSD-TLPTIHVVVTPYTSRVPQKAELTRMANT 121
QY 105 FRQVAQLHWILVEDAAARSELVSRFLARAGLPSTHLHVPTPRYK-----RPGLPRATE 158

```

```
Db 122 LLHVPNLHVLVVEDAPRRPTLTABLLRDTGLNLYTHLHVETPRNYKLRGDARDPRIPRGTM 181
Qy 159 ORNAGLAWLRQRHQHQAQPGVLPFADDDNTYSLELEQEMETTRKVSVPVGLVGGRYE 218
Db 182 ORNLALRWLRTRFPNSTQPGVVVFADDDNTYSLELEFEEMKSTRVSVPVAFVGGRLYE 241
Qy 219 RPLVEN-GKVVGVWYTWGRADRPFAIDMAGFAVSLQVILSNPKAVFKRGSGPQGMQESDFL 277
Db 242 APRVNGAGKVVGMKTVDPHPFPFALDMAGFAVNLRLILQSRQAYFKLRGVKGQYQESSLL 301
Qy 278 KQITTVLELEPKANNCTKVLVWHTRTKVKNLANEPKVKHLDTVKIEV 323
Db 302 RELVTLNDLEPKAANCTKILVWHTRTKPKVPLVNEGKGFOTDPSVEI 347
```

## RESULT 5

```
US-09-270-767-33192
; Sequence 33192, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 33192
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-33192
```

```
Query Match 28.5%; Score 486.5; DB 2; Length 291;
Best Local Similarity 40.7%; Pred. No. 1.5e-43;
Matches 103; Conservative 45; Mismatches 90; Indels 15; Gaps 7;
```

```
Qy 68 QSRPQPPEPOLPIYAITPTYSRPVQKAEITRLANTFRQVQALHWTILVEDAAARSELVS 127
Db 2 QDKPQ-SDYVQLPVIYFTPTYPREQIPETRLAHTLLHIPRLHVLVADDEKCNMDYMD 60
Qy 128 RFLARAGLPSTHLHVPTRRYK--RPGLPRAEQRNAGLAWLRQRHQHQAQPGVLPFAD 185
Db 61 TLLYRFGMPFTHMVSMPKSRNEKPA-PRGVANRAALQWIRQHN---LTNGILYFGD 115
Qy 186 DDNTYSLELEFOEMTTRKVSVPVGLVGGRRYERPLVNGKVGWYTGWRADRPFAIDMA 245
Db 116 DDNTYDLRLFSEIRKQTVSMFPVGLIADYGVSGPVVRKGVAFVFLDSWVAGRRWPVDMA 175
Qy 246 GFAVSLQVILSNPKAVFKRGSGPQGMQESDFLQOI-TTVELEPKANNCTKVLVWHTRTE 304
Db 176 GFAVNLEYMAQYP---YVNMPYKPGYEDDLFLRSIGLQMNLIIEPRGNCTEILVWHTQTK 232
Qy 305 KVNLAN---EPKY 314
Db 233 SKKLGWVRLESKY 245
```

## RESULT 6

```
US-09-270-767-48409
; Sequence 48409, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48409
; LENGTH: 291
```

```
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-48409
```

```
Query Match 28.5%; Score 486.5; DB 2; Length 291;
Best Local Similarity 40.7%; Pred. No. 1.5e-43;
Matches 103; Conservative 45; Mismatches 90; Indels 15; Gaps 7;
```

```
Qy 68 QSRPQPPEPOLPIYAITPTYSRPVQKAEITRLANTFRQVQALHWTILVEDAAARSELVS 127
Db 2 QDKPQ-SDYVQLPVIYFTPTYPREQIPETRLAHTLLHIPRLHVLVADDEKCNMDYMD 60
Qy 128 RFLARAGLPSTHLHVPTRRYK--RPGLPRAEQRNAGLAWLRQRHQHQAQPGVLPFAD 185
Db 61 TLLYRFGMPFTHMVSMPKSRNEKPA-PRGVANRAALQWIRQHN---LTNGILYFGD 115
Qy 186 DDNTYSLELEFOEMTTRKVSVPVGLVGGRRYERPLVNGKVGWYTGWRADRPFAIDMA 245
Db 116 DDNTYDLRLFSEIRKQTVSMFPVGLIADYGVSGPVVRKGVAFVFLDSWVAGRRWPVDMA 175
Qy 246 GFAVSLQVILSNPKAVFKRGSGPQGMQESDFLQOI-TTVELEPKANNCTKVLVWHTRTE 304
Db 176 GFAVNLEYMAQYP---YVNMPYKPGYEDDLFLRSIGLQMNLIIEPRGNCTEILVWHTQTK 232
Qy 305 KVNLAN---EPKY 314
Db 233 SKKLGWVRLESKY 245
```

## RESULT 7

```
US-09-270-767-45263
; Sequence 45263, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45263
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-45263
```

```
Query Match 13.8%; Score 236; DB 2; Length 142;
Best Local Similarity 35.3%; Pred. No. 3.4e-17;
Matches 48; Conservative 29; Mismatches 45; Indels 14; Gaps 4;
```

```
Qy 197 EMETTRKVSVPVGLVGGRRYERPLVNGKVGWYTGWRADRPFAIDMAGFAVSLQVILS 256
Db 8 QMRYISKVAMVPVGLVTKTVGVSPIIQAGKLVGYDGMGKRYKYPVDMAGFAVSVKFLKE 67
Qy 257 NPKAVFKRGSGPQGMQESDFLQOITVVE--ELEPKANNCTKVLVWHTRTEK---VNLANE 311
Db 68 RPNA---QMPFKPGYEDGFLRSLAPLDADAEIELLADECRDILTWTHTQTKNPAQAQLNR 124
Qy 312 PKY-----HLDTVKI 321
Db 125 TRYKNTNLEHIDRLLV 140
```

## RESULT 8

```
US-09-252-991A-18119
; Sequence 18119, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
```

```

; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18119
; LENGTH: 521
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18119

```

Query Match	7.7%;	Score 130.5;	DB 2;	Length 521;
Best Local Similarity	30.2%;	Pred. No. 4.7e-05;		
Matches	55;	Conservative	18;	Mismatches 70; Indels 39; Gaps 8;
Qy	24	DVDTRRVVPPLTRPPYFSPYAVRGGAARLPLRRGGCPAHGQTKRNQSR-----POQPQEPQ	78	
Db	56	DRDRRQPRRLCP-----APRGPRGA-----RRGRPDGPTRRRQALLAHLPRHPADH	106	
Qy	79	LPTIYALTPTSPYRPVQAELTRLANTRFVAQLHWILVEDAAARSELVSRFLARAGLPST	138	
Db	107	LPQRPAAPDPGPRGQAHGGAQQQPRGAAR-----RAEEAVSRTEVRRIRCCSRGPVA	162	
Qy	139	H-----LHVPTPRRYKRPQ-----LP-----RATEQNAAGLAWLRQHQQHQAQ-----	177	
Db	163	HGRRRRHRPDPGRLQRTGDEPGVLPQRPRLRLARSFRAGLGLAGRRRQPDERGQVPR	222	
Qy	178	PG	179	
Db	223	PG	224	

```

RESULT 9
US-09-252-991A-25264
; Sequence 25264, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25264
; LENGTH: 603
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25264

```

Query Match	6.8%;	Score 115.5;	DB 2;	Length 603;
Best Local Similarity	27.1%;	Pred. No. 0.0024;		
Matches 55;	Conservative 14;	Mismatches 59;	Indels 75;	Gaps 11;
QY	26	DTRRPVPLTPRPVSPVAVGRG-----GARLPLRRGGPAHGQKKNRSRQPQPQPQL	79	
Db	27	DSRAPPAPIPHPAFLTA-GRGRQLHQLVFLRRHP--GTARTATQPAFLVP--	81	
QY	80	PT-----IYAITYSRPVQKAELTLANTFRQVAQLHWLV	116	
Db	82	PTGALRPPLAQRPRGDRRQPDPAARSGLRPRAAPPAAGAYPRRGAQPA-----	133	
QY	117	EDAAARSELYSRLARAGLSTHLHP--TPRR-----YKPGGLPRATEQRNAG	163	
Db	134	--PAARLITAPPVATAGVPTRNQHEPFRPORADRRRLSRQPARRSPGLP-----QRRTG	187	
QY	164	LAWLRQRH-----QHORA	176	

```

Db      188  ----RRRHPAGVGPORLPQHRRKA 206
      ||:|:|
      ||:|:|
      ||:|:|

RESULT 10
US-09-050-739-70
; Sequence 70, Application US/09050739
; Patent No. 6641814
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: NIELSEN, Rikke
; APPLICANT: OETTINGER, Thomas
; APPLICANT: RASMUSSEN, Peter Birk
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: WELDKINGH, Karin
; APPLICANT: FLORIO, Walter
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS
; TITLE OF INVENTION: DERIVED FROM M. TUBERCULOSIS
; FILE REFERENCE: 670001-2002.1
; CURRENT APPLICATION NUMBER: US/09/050,
; CURRENT FILING DATE: 1998-03-30
; EARLIER APPLICATION NUMBER: 0376/97
; EARLIER FILING DATE: 1997-04-02
; EARLIER APPLICATION NUMBER: 1277/97
; EARLIER FILING DATE: 1997-11-10
; EARLIER APPLICATION NUMBER: 60/044,624
; EARLIER FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/070,488
; EARLIER FILING DATE: 1998-01-05
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 70
; LENGTH: 666
; TYPE: PRN
; ORGANISM: Mycobacterium tuberculosis
US-09-050-739-70

```

Query Match	6.5%; Score 111; DB 2; Length 566;
Best Local Similarity	21.3%; Pred. No. 0.0084;
Matches	73; Conservative 28; Mismatches 120; Indels 122; Gaps 13;
Qy	30 PVPELTTPRPFSYAVGCGARLPLR-----RGG-----PAHQTKQRNQRSP 71
Db	129 PKPPTPMPITAGPAPTTESQLAPRPPTQTPTGAPQOQSPAPHVPSHGPHQPRRTAP 188
Qy	72 -----QPOPEPLQTIYAI-----TPTYSRPVQKAELTRLANTRFQVLAQHWIL 115
Db	189 APWAKMPIGEPAPPAPSRPSASPAEPTRPAPQHSRRARRGHRV-TDTERNVGKVATGP 247
Qy	116 VEDAAARSELVSRFLARAGLPSTHLHVTPRRYKRPGL-----PPATQRNAGLAWL 167
Db	248 SIQARLRAEASGAQLAPGTGPEPSAPLQGPQSRSYLAPTRPAPTEPPSPSPQRNSGRRAE 307
Qy	168 RQRH-----QHQRAPQGVLFADDDNTYSLELFOEMRTTRKVSVPVLVGRRVERPLV 222
Db	308 RRVHPDLAAQHAQAQPSDITAA-----TT-----GRRRKRA-- 339
Qy	223 ENGVGVWYGTWRADRPFAIDMAGFAVSLQVILSNPKAVFKRQSGPQMGBSDFLKQIT 282
Db	340 -----APDLDAQSLRPAAGPK-----VKK 361
Qy	283 VEELEPKANNCTKVLV-----WHRTREKVNLANEP--KYHLD 317
Db	362 VKPQKPRATPKPKVVSORGWHRWVHALTRINLGLSPDEKVELD 404

```

RESULT 11
US-09-252-991A-26907
; Sequence 26907, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND

```



; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 23212  
; LENGTH: 720  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-23212

Query Match 6.2%; Score 106.5; DB 2; Length 720;  
Best Local Similarity 27.7%; Pred. No. 0.029;  
Matches 53; Conservative 16; Mismatches 69; Indels 53; Gaps 9;  
Qy 29 RVPVPLTP--RPYFSPYA-----VGRGARLPLRRGGPAHGTQKRNQSRPQPEPQL 79  
Db 121 RPAPRRLPALRPY--PGAARGARRTAGAG--RQPLAGAMGAGAQAHR-----ANFGPAL 171  
Qy 80 PT-----IYAITPTYSRP-----VQKAELTRLANTFRQVAQLHLV-----116  
Db 172 PARTGHDLRMAPTPALPRGHRGAGSRQGDPAADRLRPLQQPLGVHLDPVPPAGRHHTGTI 231  
Qy 117 -----EDAAARSELVSRFLARAGLPSTHLHVPTPRYKRPGLPRATEQRNAGLAWL 167  
Db 232 PQCQPRPGHADATVRACVAAMLAATIRSSHANRQPPSEHPRGVPRLPQ---AGLQRL 288  
Qy 168 RQRHQHQAQP 178  
Db 289 RRAHRPPRLFP 299

RESULT 15  
US-09-252-991A-17917  
; Sequence 17917, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 17917  
; LENGTH: 388  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-17917

Query Match 6.2%; Score 105.5; DB 2; Length 388;  
Best Local Similarity 28.0%; Pred. No. 0.014;  
Matches 45; Conservative 16; Mismatches 61; Indels 39; Gaps 8;  
Qy 20 IIMLDVDTTRPVPPLTPRPYPFVYVGRGARLPLRRGGPAHGTQKRNQSRPQPEPQL 79  
Db 80 LVNIRLHEHTPDPAPATWRA-AVRGCGPRIQRRTG-SHGCRQAQTRLPGRP--135  
Qy 80 PTIYAITPTYSRVQKAELTRLANTFRQVAQLHLVVEDAAARSELVSR-FLARAGLPST 138  
Db 136 -----AAAAGAARPDRR-----VFRQSSQAP--AADQPVARARLFQRHQHQTAGSLGT 181  
Qy 139 HLHVPTPRYKRGRLPRATEQRNAGLAWLQRHQHQAQP 179  
Db 182 G---PRPRPCARPG-----PGDHRRVEPG 202

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2006, 10:55:28 ; Search time 165 Seconds  
(without alignments)  
817.933 Million cell updates/sec

Title: US-10-634-905-2

Perfect score: 1705

Sequence: 1 MKSALFTRFFILLPWILVI.....EKNVLANEPKYHLDTVKIEV 323

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

- Database : Published Applications AA\_Main:\*
- 1: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pgp:\*
  - 2: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pgp:\*
  - 3: /cgn2\_6/prodata/1/pubpaa/US09\_PUBCOMB.pgp:\*
  - 4: /cgn2\_6/prodata/1/pubpaa/US10A\_PUBCOMB.pgp:\*
  - 5: /cgn2\_6/prodata/1/pubpaa/US10B\_PUBCOMB.pgp:\*
  - 6: /cgn2\_6/prodata/1/pubpaa/US11\_PUBCOMB.pgp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1705	100.0	323	3	US-09-816-095-2
2	1705	100.0	323	4	US-10-634-905-2
3	1526.5	89.5	324	3	US-09-816-095-4
4	1526.5	89.5	324	4	US-10-634-905-4
5	1039	60.9	197	3	US-09-816-095-5
6	1039	60.9	197	4	US-10-634-905-5
7	653	38.3	335	5	US-10-698-190-16
8	653	38.3	335	6	US-10-723-860-4238
9	502.5	29.5	479	6	US-11-097-143-13308
10	487.5	28.6	443	6	US-11-097-143-7506
11	422	24.8	145	4	US-10-080-254-93
12	422	24.8	145	4	US-10-242-355-637
13	406	23.8	691	5	US-10-450-763-30691
14	324.5	19.3	201	4	US-10-403-571-34
15	312.5	18.0	351	4	US-10-437-963-174074
16	305.5	17.9	331	4	US-10-425-115-330485
17	293	17.2	414	4	US-10-424-599-147426
18	256.5	15.0	446	4	US-10-437-963-135854
19	249	14.6	448	4	US-10-425-115-366742
20	246	14.4	451	4	US-10-425-115-292707
21	246	14.4	479	4	US-10-425-114-68877
22	244.5	14.3	528	4	US-10-425-115-360409
23	241.5	14.2	460	4	US-10-425-114-62647
24	229	13.4	524	4	US-10-437-963-177911
25	227	13.3	411	4	US-10-425-114-54143
26	221	13.0	385	4	US-10-425-115-345009
27	213	12.5	445	4	US-10-437-963-107957

28	210.5	12.3	342	4	US-10-424-599-178009	Sequence 178009,
29	209.5	12.3	316	4	US-10-425-114-53567	Sequence 53567, A
30	202.5	11.9	344	4	US-10-424-599-217815	Sequence 217815,
31	202.5	11.9	471	4	US-10-437-963-125866	Sequence 125866,
32	201.5	11.8	494	4	US-10-424-599-233474	Sequence 233474,
33	200	11.7	371	4	US-10-437-963-200415	Sequence 200415,
34	194.5	11.4	167	5	US-10-450-763-43912	Sequence 43912, A
35	188.5	11.1	367	4	US-10-425-115-290232	Sequence 290232,
36	188	11.0	549	4	US-10-437-963-102770	Sequence 102770,
37	181	10.6	235	4	US-10-437-963-140001	Sequence 140001,
38	179	10.5	883	4	US-10-425-115-296340	Sequence 296340,
39	171	10.0	273	4	US-10-425-114-50568	Sequence 50568, A
40	162	9.5	384	4	US-10-425-115-350198	Sequence 350198,
41	145	8.5	156	4	US-10-767-701-48703	Sequence 48703, A
42	141.5	8.3	378	4	US-10-425-114-42131	Sequence 42131, A
43	139	8.2	40	4	US-10-029-386-33145	Sequence 33145, A
44	138.5	8.1	374	4	US-10-425-115-364687	Sequence 364687,
45	137.5	8.1	267	4	US-10-425-114-54605	Sequence 54605, A

ALIGNMENTS

RESULT 1

US-09-816-095-2  
; Sequence 2, Application US/09816095  
; Patent No. US20020137164A1  
; GENERAL INFORMATION:  
; APPLICANT: GAN, Weiniu  
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES  
; FILE REFERENCE: CL001147  
; CURRENT APPLICATION NUMBER: US/09/816,095  
; CURRENT FILING DATE: 2001-03-26  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 323  
; TYPE: PRT  
; ORGANISM: Human  
US-09-816-095-2

Query Match 100.0%; Score 1705; DB 3; Length 323;  
Best Local Similarity 100.0%; Pred. No. 2.9e-149;  
Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MKSALFTRFFILLPWILVIIMLDVDTTRRPVPLTPRPYFSPYAVGCGARLPLRRGGA	60
Db	1	MKSALFTRFFILLPWILVIIMLDVDTTRRPVPLTPRPYFSPYAVGCGARLPLRRGGA	60
Qy	61	HGTQKNSRQPOPEPOLPTIYAITPTYSRPVQKAELTRANTFRQVQLHWILVEDAA	120
Db	61	HGTQKNSRQPOPEPOLPTIYAITPTYSRPVQKAELTRANTFRQVQLHWILVEDAA	120
Qy	121	AKSELVSRFLAAGLPSLTHLVPTPRYKRPGLPRATEORNAGLWLRQHQHQAQPGV	180
Db	121	AKSELVSRFLAAGLPSLTHLVPTPRYKRPGLPRATEORNAGLWLRQHQHQAQPGV	180
Qy	181	LFFFADDDNTYSLELFQEMRTTRKVSVPVGLVGGRRYERPLVGVKGYTGWADRPFF	240
Db	181	LFFFADDDNTYSLELFQEMRTTRKVSVPVGLVGGRRYERPLVGVKGYTGWADRPFF	240
Qy	241	AIDMAGFVSLQVILSNPKAVFKRGSGQMGSDFLKQITTTVEELPKANNCTKVLVWH	300
Db	241	AIDMAGFVSLQVILSNPKAVFKRGSGQMGSDFLKQITTTVEELPKANNCTKVLVWH	300
Qy	301	TRTEKVNLANEPKYHLDTVKIEV	323
Db	301	TRTEKVNLANEPKYHLDTVKIEV	323

RESULT 2

US-10-634-905-2		US-10-634-905-2	
;	Sequence 2, Application US/10634905	;	Sequence 2, Application US/10634905
;	Publication No. US20040067225A1	;	Publication No. US20040067225A1
;	GENERAL INFORMATION:	;	GENERAL INFORMATION:
;	APPLICANT: GAN, Weinliu et al.	;	APPLICANT: GAN, Weinliu et al.
;	TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC	;	TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
;	TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES	;	TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
;	TITLE OF INVENTION: THEREOF	;	TITLE OF INVENTION: THEREOF
;	FILE REFERENCE: CL001147-DIV	;	FILE REFERENCE: CL001147-DIV
;	CURRENT APPLICATION NUMBER: US/10/634,905	;	CURRENT APPLICATION NUMBER: US/10/634,905
;	CURRENT FILING DATE: 2003-08-06	;	CURRENT FILING DATE: 2003-08-06
;	NUMBER OF SEQ ID NOS: 5	;	NUMBER OF SEQ ID NOS: 5
;	SOFTWARE: FastSeq for Windows Version 4.0	;	SOFTWARE: FastSeq for Windows Version 4.0
;	SEQ ID NO 2	;	SEQ ID NO 2
;	LENGTH: 323	;	LENGTH: 323
;	TYPE: PRT	;	TYPE: PRT
;	ORGANISM: Homo sapiens	;	ORGANISM: Homo sapiens
US-10-634-905-2		US-10-634-905-2	
Query Match 100.0%; Score 1705; DB 4; Length 323;		Query Match 100.0%; Score 1705; DB 4; Length 323;	
Best Local Similarity 100.0%; Pred. No. 2.9e-149;		Best Local Similarity 100.0%; Pred. No. 2.9e-149;	
Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MKSALTRFRFFILLPWILVIMLDVDRPRPAPQLTSRPFSPHTVCGGSRVPLRRSSPG 60	QY	1 MKSALTRFRFFILLPWILVIMLDVDRPRPAPQLTSRPFSPHTVCGGSRVPLRRSSPG 60
DB	1 MKSALTRFRFFILLPWILVIMLDVDRPRPAPQLTSRPFSPHTVCGGSRVPLRRSSPG 60	DB	1 MKSALTRFRFFILLPWILVIMLDVDRPRPAPQLTSRPFSPHTVCGGSRVPLRRSSPG 60
QY	61 HGTQKNSRQRPQPEPOLPTIYAITPTYSRPVQKAELTRLANTFRQVAQLHWILVEDAA 120	QY	61 HGTQKNSRQRPQPEPOLPTIYAITPTYSRPVQKAELTRLANTFRQVAQLHWILVEDAA 120
DB	61 HGTQKNSRQRPQPEPOLPTIYAITPTYSRPVQKAELTRLANTFRQVAQLHWILVEDAA 120	DB	61 HGTQKNSRQRPQPEPOLPTIYAITPTYSRPVQKAELTRLANTFRQVAQLHWILVEDAA 120
QY	121 ARSELVSRFLARAGLPSTHLHVPTRPRYKRPGLPRATEQORNAGLAWLRQHQHQAQPGV 180	QY	121 ARSELVSRFLARAGLPSTHLHVPTRPRYKRPGLPRATEQORNAGLAWLRQHQHQAQPGV 180
DB	121 ARSELVSRFLARAGLPSTHLHVPTRPRYKRPGLPRATEQORNAGLAWLRQHQHQAQPGV 180	DB	121 ARSELVSRFLARAGLPSTHLHVPTRPRYKRPGLPRATEQORNAGLAWLRQHQHQAQPGV 180
QY	181 LFFADDDNTYSLELFOEMRTTRKVSVMVGLVGGRRYERPLVGVGKVGWYTGWRADRP 240	QY	181 LFFADDDNTYSLELFOEMRTTRKVSVMVGLVGGRRYERPLVGVGKVGWYTGWRADRP 240
DB	181 LFFADDDNTYSLELFOEMRTTRKVSVMVGLVGGRRYERPLVGVGKVGWYTGWRADRP 240	DB	181 LFFADDDNTYSLELFOEMRTTRKVSVMVGLVGGRRYERPLVGVGKVGWYTGWRADRP 240
QY	241 AIDMAGFAVSLQVILSNPKAVFKRRGSGQGMQESDFLKQITTVDELEPKANNCTKVLVW 300	QY	241 AIDMAGFAVSLQVILSNPKAVFKRRGSGQGMQESDFLKQITTVDELEPKANNCTKVLVW 300
DB	241 AIDMAGFAVSLQVILSNPKAVFKRRGSGQGMQESDFLKQITTVDELEPKANNCTKVLVW 300	DB	241 AIDMAGFAVSLQVILSNPKAVFKRRGSGQGMQESDFLKQITTVDELEPKANNCTKVLVW 300
QY	301 TRTEKVNLANEPKYHLDTVKIEV 323	QY	301 TRTEKVNLANEPKYHLDTVKIEV 323
DB	301 TRTEKVNLANEPKYHLDTVKIEV 323	DB	301 TRTEKVNLANEPKYHLDTVKIEV 323
RESULT 3		RESULT 3	
US-09-816-095-4		US-09-816-095-4	
;	Sequence 4, Application US/09816095	;	Sequence 4, Application US/09816095
;	Patent No. US20020137164A1	;	Patent No. US20020137164A1
;	GENERAL INFORMATION:	;	GENERAL INFORMATION:
;	APPLICANT: GAN, Weinliu	;	APPLICANT: GAN, Weinliu
;	TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC	;	TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
;	TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES	;	TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
;	TITLE OF INVENTION: THEREOF	;	TITLE OF INVENTION: THEREOF
;	FILE REFERENCE: CL001147	;	FILE REFERENCE: CL001147
;	CURRENT APPLICATION NUMBER: US/09/816,095	;	CURRENT APPLICATION NUMBER: US/09/816,095
;	CURRENT FILING DATE: 2001-03-26	;	CURRENT FILING DATE: 2001-03-26
;	NUMBER OF SEQ ID NOS: 5	;	NUMBER OF SEQ ID NOS: 5
;	SOFTWARE: FastSeq for Windows Version 4.0	;	SOFTWARE: FastSeq for Windows Version 4.0
;	SEQ ID NO 4	;	SEQ ID NO 4
;	LENGTH: 324	;	LENGTH: 324
;	TYPE: PRT	;	TYPE: PRT
;	ORGANISM: Rattus norvegicus	;	ORGANISM: Rattus norvegicus
US-09-816-095-4		US-09-816-095-4	
Query Match 89.5%; Score 1526.5; DB 3; Length 324;		Query Match 89.5%; Score 1526.5; DB 3; Length 324;	
Best Local Similarity 89.8%; Pred. No. 1e-132;		Best Local Similarity 89.8%; Pred. No. 1e-132;	
Matches 291; Conservative 10; Mismatches 22; Indels 1; Gaps 1;		Matches 291; Conservative 10; Mismatches 22; Indels 1; Gaps 1;	
QY	1 MKSALTRFRFFILLPWILVIMLDVDRPRPAPQLTSRPFSPHTVCGGSRVPLRRSSPG 60	QY	1 MKSALTRFRFFILLPWILVIMLDVDRPRPAPQLTSRPFSPHTVCGGSRVPLRRSSPG 60
DB	1 MKSALTRFRFFILLPWILVIMLDVDRPRPAPQLTSRPFSPHTVCGGSRVPLRRSSPG 60	DB	1 MKSALTRFRFFILLPWILVIMLDVDRPRPAPQLTSRPFSPHTVCGGSRVPLRRSSPG 60
QY	61 HGTQKNSRQRPQPEPOLPTIYAITPTYSRPVQKAELTRLANTFRQVAQLHWILVEDAA 119	QY	61 HGTQKNSRQRPQPEPOLPTIYAITPTYSRPVQKAELTRLANTFRQVAQLHWILVEDAA 119
DB	61 HGTQKNSRQRPQPEPOLPTIYAITPTYSRPVQKAELTRLANTFRQVAQLHWILVEDAA 119	DB	61 HGTQKNSRQRPQPEPOLPTIYAITPTYSRPVQKAELTRLANTFRQVAQLHWILVEDAA 119
QY	120 AARSELVSRFLARAGLPSTHLHVPTRPRYKRPGLPRATEQORNAGLAWLRQHQHQAQPG 179	QY	120 AARSELVSRFLARAGLPSTHLHVPTRPRYKRPGLPRATEQORNAGLAWLRQHQHQAQPG 179
DB	120 AARSELVSRFLARAGLPSTHLHVPTRPRYKRPGLPRATEQORNAGLAWLRQHQHQAQPG 179	DB	120 AARSELVSRFLARAGLPSTHLHVPTRPRYKRPGLPRATEQORNAGLAWLRQHQHQAQPG 179
QY	180 VLFADDDNTYSLELFOEMRTTRKVSVMVGLVGGRRYERPLVGVGKVGWYTGWRADRP 239	QY	180 VLFADDDNTYSLELFOEMRTTRKVSVMVGLVGGRRYERPLVGVGKVGWYTGWRADRP 239
DB	180 VLFADDDNTYSLELFOEMRTTRKVSVMVGLVGGRRYERPLVGVGKVGWYTGWRADRP 239	DB	180 VLFADDDNTYSLELFOEMRTTRKVSVMVGLVGGRRYERPLVGVGKVGWYTGWRADRP 239
QY	240 PAIDMAGFAVSLQVILSNPKAVFKRRGSGQGMQESDFLKQITTVDELEPKANNCTKVLVW 299	QY	240 PAIDMAGFAVSLQVILSNPKAVFKRRGSGQGMQESDFLKQITTVDELEPKANNCTKVLVW 299
DB	240 PAIDMAGFAVSLQVILSNPKAVFKRRGSGQGMQESDFLKQITTVDELEPKANNCTKVLVW 299	DB	240 PAIDMAGFAVSLQVILSNPKAVFKRRGSGQGMQESDFLKQITTVDELEPKANNCTKVLVW 299
QY	300 HTRTEKVNLANEPKYHLDTVKIEV 323	QY	300 HTRTEKVNLANEPKYHLDTVKIEV 323
DB	300 HTRTEKVNLANEPKYHLDTVKIEV 323	DB	300 HTRTEKVNLANEPKYHLDTVKIEV 323
US-10-634-905-4		US-10-634-905-4	
;	Sequence 4, Application US/10634905	;	Sequence 4, Application US/10634905
;	Publication No. US20040067225A1	;	Publication No. US20040067225A1
;	GENERAL INFORMATION:	;	GENERAL INFORMATION:
;	APPLICANT: GAN, Weinliu et al.	;	APPLICANT: GAN, Weinliu et al.
;	TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC	;	TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
;	TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES	;	TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
;	TITLE OF INVENTION: THEREOF	;	TITLE OF INVENTION: THEREOF
;	FILE REFERENCE: CL001147-DIV	;	FILE REFERENCE: CL001147-DIV
;	CURRENT APPLICATION NUMBER: US/10/634,905	;	CURRENT APPLICATION NUMBER: US/10/634,905
;	CURRENT FILING DATE: 2003-08-06	;	CURRENT FILING DATE: 2003-08-06
;	NUMBER OF SEQ ID NOS: 5	;	NUMBER OF SEQ ID NOS: 5
;	SOFTWARE: FastSeq for Windows Version 4.0	;	SOFTWARE: FastSeq for Windows Version 4.0
;	SEQ ID NO 4	;	SEQ ID NO 4
;	LENGTH: 324	;	LENGTH: 324
;	TYPE: PRT	;	TYPE: PRT
;	ORGANISM: Rattus norvegicus	;	ORGANISM: Rattus norvegicus
US-10-634-905-4		US-10-634-905-4	
Query Match 89.5%; Score 1526.5; DB 4; Length 324;		Query Match 89.5%; Score 1526.5; DB 4; Length 324;	
Best Local Similarity 89.8%; Pred. No. 1e-132;		Best Local Similarity 89.8%; Pred. No. 1e-132;	
Matches 291; Conservative 10; Mismatches 22; Indels 1; Gaps 1;		Matches 291; Conservative 10; Mismatches 22; Indels 1; Gaps 1;	
QY	1 MKSALTRFRFFILLPWILVIMLDVDRPRPAPQLTSRPFSPHTVCGGSRVPLRRSSPG 60	QY	1 MKSALTRFRFFILLPWILVIMLDVDRPRPAPQLTSRPFSPHTVCGGSRVPLRRSSPG 60
DB	1 MKSALTRFRFFILLPWILVIMLDVDRPRPAPQLTSRPFSPHTVCGGSRVPLRRSSPG 60	DB	1 MKSALTRFRFFILLPWILVIMLDVDRPRPAPQLTSRPFSPHTVCGGSRVPLRRSSPG 60
QY	61 HGTQKNSRQRPQPEPOLPTIYAITPTYSRPVQKAELTRLANTFRQVAQLHWILVEDAA 119	QY	61 HGTQKNSRQRPQPEPOLPTIYAITPTYSRPVQKAELTRLANTFRQVAQLHWILVEDAA 119
DB	61 HGTQKNSRQRPQPEPOLPTIYAITPTYSRPVQKAELTRLANTFRQVAQLHWILVEDAA 119	DB	61 HGTQKNSRQRPQPEPOLPTIYAITPTYSRPVQKAELTRLANTFRQVAQLHWILVEDAA 119
QY	120 AARSELVSRFLARAGLPSTHLHVPTRPRYKRPGLPRATEQORNAGLAWLRQHQHQAQPG 179	QY	120 AARSELVSRFLARAGLPSTHLHVPTRPRYKRPGLPRATEQORNAGLAWLRQHQHQAQPG 179
DB	120 AARSELVSRFLARAGLPSTHLHVPTRPRYKRPGLPRATEQORNAGLAWLRQHQHQAQPG 179	DB	120 AARSELVSRFLARAGLPSTHLHVPTRPRYKRPGLPRATEQORNAGLAWLRQHQHQAQPG 179
QY	180 VLFADDDNTYSLELFOEMRTTRKVSVMVGLVGGRRYERPLVGVGKVGWYTGWRADRP 239	QY	180 VLFADDDNTYSLELFOEMRTTRKVSVMVGLVGGRRYERPLVGVGKVGWYTGWRADRP 239
DB	180 VLFADDDNTYSLELFOEMRTTRKVSVMVGLVGGRRYERPLVGVGKVGWYTGWRADRP 239	DB	180 VLFADDDNTYSLELFOEMRTTRKVSVMVGLVGGRRYERPLVGVGKVGWYTGWRADRP 239
QY	240 PAIDMAGFAVSLQVILSNPKAVFKRRGSGQGMQESDFLKQITTVDELEPKANNCTKVLVW 299	QY	240 PAIDMAGFAVSLQVILSNPKAVFKRRGSGQGMQESDFLKQITTVDELEPKANNCTKVLVW 299
DB	240 PAIDMAGFAVSLQVILSNPKAVFKRRGSGQGMQESDFLKQITTVDELEPKANNCTKVLVW 299	DB	240 PAIDMAGFAVSLQVILSNPKAVFKRRGSGQGMQESDFLKQITTVDELEPKANNCTKVLVW 299
QY	300 HTRTEKVNLANEPKYHLDTVKIEV 323	QY	300 HTRTEKVNLANEPKYHLDTVKIEV 323
DB	300 HTRTEKVNLANEPKYHLDTVKIEV 323	DB	300 HTRTEKVNLANEPKYHLDTVKIEV 323

```
RESULT 5
US-09-816-095-5
; Sequence 5, Application US/09816095
; Patent No. US20020137164A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weinuu
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001147
; CURRENT APPLICATION NUMBER: US/09/816,095
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Human
US-09-816-095-5

Query Match          60.9%; Score 1039; DB 3; Length 197;
Best Local Similarity 100.0%; Pred. No. 7.7e-88;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKSALFTRFFILLPWILVIMLDVTRRPVPLTPRPYSPYAVGRGGARLPLRRGGPA 60
Db 1 MKSALFTRFFILLPWILVIMLDVTRRPVPLTPRPYSPYAVGRGGARLPLRRGGPA 60

Qy 61 HGQKRNQSRPQPOPEQLPTIYAITYSRPVQKAELTRANTFRVAQLHWILVEDAA 120
Db 61 HGQKRNQSRPQPOPEQLPTIYAITYSRPVQKAELTRANTFRVAQLHWILVEDAA 120

Qy 121 ARSELVSRFLARAGLPSTHLHVTPPRYKRPGLPRATEQRNAGLAWLRQRHQHQAQPGV 180
Db 121 ARSELVSRFLARAGLPSTHLHVTPPRYKRPGLPRATEQRNAGLAWLRQRHQHQAQPGV 180

Qy 181 LFFADDDNTYSLELFQE 197
Db 181 LFFADDDNTYSLELFQE 197

RESULT 6
US-10-634-905-5
; Sequence 5, Application US/10634905
; Publication No. US20040067225A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weinuu et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001147-DIV
; CURRENT APPLICATION NUMBER: US/10/634,905
; CURRENT FILING DATE: 2003-08-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-634-905-5

Query Match          60.9%; Score 1039; DB 4; Length 197;
Best Local Similarity 100.0%; Pred. No. 7.7e-88;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKSALFTRFFILLPWILVIMLDVTRRPVPLTPRPYSPYAVGRGGARLPLRRGGPA 60
Db 1 MKSALFTRFFILLPWILVIMLDVTRRPVPLTPRPYSPYAVGRGGARLPLRRGGPA 60

Qy 61 HGQKRNQSRPQPOPEQLPTIYAITYSRPVQKAELTRANTFRVAQLHWILVEDAA 120
Db 61 HGQKRNQSRPQPOPEQLPTIYAITYSRPVQKAELTRANTFRVAQLHWILVEDAA 120
```

```
Qy 121 ARSELVSRFLARAGLPSTHLHVTPPRYKRPGLPRATEQRNAGLAWLRQRHQHQAQPGV 180
Db 121 ARSELVSRFLARAGLPSTHLHVTPPRYKRPGLPRATEQRNAGLAWLRQRHQHQAQPGV 180

Qy 181 LFFADDDNTYSLELFQE 197
Db 181 LFFADDDNTYSLELFQE 197

RESULT 7
US-10-698-190-16
; Sequence 16, Application US/10698190
; Publication No. US20040147469A1
; GENERAL INFORMATION:
; APPLICANT: Silver, J.
; APPLICANT: Grimpe, B.
; TITLE OF INVENTION: Methods of Inhibiting Glial Scar Formation
; FILE REFERENCE: CMRU-P01-018
; CURRENT APPLICATION NUMBER: US/10/698,190
; CURRENT FILING DATE: 2003-10-31
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-698-190-16

Query Match          38.3%; Score 653; DB 4; Length 335;
Best Local Similarity 43.9%; Pred. No. 9e-52;
Matches 150; Conservative 41; Mismatches 93; Indels 58; Gaps 10;

Qy 1 MKSALFTRFPILLPWILVIMLDVTRRRP---VPPLTPRPYSPYAVGRGGA----- 50
Db 5 LKNVFLAYFLVSLAGLYALVQLG-----QPCDCLPPL-----RAAAEQLRQKD 48

Qy 51 -----RLPLRRGGPAHGTQKRNQSRPQPOPEQLPTIYAITYSRPVQKAELTRAN 103
Db 49 LRISQLQAELELRPPPA-----PAQPEPEALPTIVVVTPTVARLVQKAEVLVSQ 98

Qy 104 TFRQVAQLHWILVEDAAARSELVSRFLARAGLPSTHLHVTPPR-----RYKRPGL--PRAT 157
Db 99 TLSLVPLRLHWLLVEDAEGPTPLVSLGAAAGLLFTHLVLTTPKAQRLEGEPEGVHPRGV 158

Qy 158 EQRNAGLAWLR-----QRHQHQAQPGVLFADDDNTYSLELFQEMETTRKVSVPV 209
Db 159 EQRNKALDWLRGGGAVGGEKDPPTGQGVVTFADDDNTYSLELFEEEMRWTRGVSVMPV 218

Qy 210 GLVGGRRYERPLVENGKVVGWYTGWRADRPFAIDMAGFAVSLQVILSNPKAVFKRRGSQP 269
Db 219 GLVGGRLRFGCPQVQDGRVVGFTAWEPSRPFVDMAGFAVALPLLLDKPNQAFDSTAPR- 277

Qy 270 GMOESDFLKQITTVBELPKANNCTKVLVWHTTEKVNLANE 311
Db 278 GHLESSLHLVDPKOLEPRAANCTRLVWHTTRTEKPKMQE 319

RESULT 8
US-10-723-860-4238
; Sequence 4238, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Nataasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
```

```
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4238
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-4238

Query Match      38.3%; Score 653; DB 5; Length 335;
Best Local Similarity 43.9%; Pred. No. 9e-52;
Matches 150; Conservative 41; Mismatches 93; Indels 58; Gaps 10;

QY 1 MKSALFRFPILLPWILIMLDVTRRP---VPPLTPRPSPYAVGGRGGA-----50
DB 5 LKNVFLAYFLVSVTAGLLYALVQLG-----QPCDCLPPL-----RAAAEQRLQKD 48
QY 51 -----RLPLRRGGPAGHTQKRNQSRPQPOPEQ-LPTIYAITPTYSRPVQKAELTRLAN 103
DB 49 LRISQLQAELELRPPPA-----PAQPPPEALPTIYVVTPTVARLVQKAELVRLSQ 98
QY 104 TFRVQAQLHWILVEDAAARSELVSRFLARAGLSTHLHVPTPR-----RYKRPGL--PRAT 157
DB 99 TLSLVPLRLHLLVEDAEGTPTLVSGLLAASGLLFTHLVLTPKAQLREGEPEGVHPRGV 158
QY 158 EQRNAGLAWLR-----QRHQHQAQPGVLPFADDDNTYSLELFOEMRTTRKVSVPV 209
DB 159 EQRNKALDWLRGGAAGVGEKDPPTGQTGVVYFADDDNTYSRELFEMRWRTGVSVPV 218
QY 210 GLVGGRRYERPLVGVGMYTGWADRPFAIDMAGFAVLSQVILSNPKAVFKRGSOP 269
DB 219 GLVGLRFGEPQVQDGRVVGFTAWEPSPFPVDMAGFAVALPFLLLDKENAPQDSTAPR- 277
QY 270 GMOESDFLKQITTVESLEPKANNCTKVLVWHTRTEKVNLANE 311
DB 278 GHLESSLLSHLVDPKOLEPRAANCTRLVNVHTRTEKPKMKQE 319

RESULT 9
US-11-097-143-13308
; Sequence 13308, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: DROSOPHILA GENES.
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13308
; LENGTH: 479
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-13308

Query Match      28.6%; Score 487.5; DB 6; Length 443;
Best Local Similarity 40.1%; Pred. No. 2.8e-36;
Matches 103; Conservative 48; Mismatches 91; Indels 15; Gaps 7;

QY 65 KENQSRPQPE-POLPTIYAITPTYSRPVQKAELTRANTFRQVAQLHWILVEDAAARS 123
DB 149 RRQFMQDQPSDYQLPVIYFVTPTYPREQLPELTRLAHTLHLPLHLWLVADQEKCN 208
QY 124 ELVSRFLARAGLPSSTHLHVPTPRRYK--RPGLPRAATEQRNAGLAWLRQHQRAQPGVL 181
DB 209 DYMDTLRYFGMPFTTHVSPWSPESKERNKPA-PRGVANRRALQWIRQHN----LTNGIL 263

Query Match      29.5%; Score 502.5; DB 6; Length 479;
Best Local Similarity 40.0%; Pred. No. 1.3e-37;
Matches 102; Conservative 46; Mismatches 86; Indels 21; Gaps 6;

QY 80 PTIYAITPTYSRPVQKAELTRANTFRQVAQLHWILVEDAAARSELVSRFLARAGLPSSTH 139
DB 231 PLYIITPTYSRPEQLAELTRGLYTLKHVVNLLVLVEDANKTNPLVGHITLDRIGVPYEX 290
QY 140 LHVPTPRRYK--RPGLPRAATEQRNAGLAWLRQHQRAQPGVLPFADDDNTYSLELFOE 197
DB 291 MVAPMPEKYKTKAKPRGVSNNRNGLEYLR-----EHATEGVLPFADDDNTYDISIFEQ 345
QY 198 MTTTRKVSVPVGLVGGRRYERPLVGVGMYTGWADRPFAIDMAGFAVLSQVILSN 257
DB 346 MRYISKVAMWVGLVTKTGVSSPIIOAGKLVGYDGMIGGRKYPVDMAGFAVSVFKLER 405
QY 258 PKAVFKRRGSGQMOESDFLKQITTVESLEPKANNCTKVLVWHTRTEK-----VNLANEP 312
DB 406 PNA---QMPFKGYSDEGFLRSFLAPLDDDAEIELLADECEDEILTWHTQTKKNAPQAALNRT 462
QY 313 KY-----HLDTVKI 321
DB 463 RYKNTNLEHIDRLLV 477

RESULT 10
US-11-097-143-7506
; Sequence 7506, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: DROSOPHILA GENES.
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7506
; LENGTH: 443
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-7506

Query Match      28.6%; Score 487.5; DB 6; Length 443;
Best Local Similarity 40.1%; Pred. No. 2.8e-36;
Matches 103; Conservative 48; Mismatches 91; Indels 15; Gaps 7;

QY 65 KENQSRPQPE-POLPTIYAITPTYSRPVQKAELTRANTFRQVAQLHWILVEDAAARS 123
DB 149 RRQFMQDQPSDYQLPVIYFVTPTYPREQLPELTRLAHTLHLPLHLWLVADQEKCN 208
QY 124 ELVSRFLARAGLPSSTHLHVPTPRRYK--RPGLPRAATEQRNAGLAWLRQHQRAQPGVL 181
DB 209 DYMDTLRYFGMPFTTHVSPWSPESKERNKPA-PRGVANRRALQWIRQHN----LTNGIL 263
```

```

; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 637
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-242-355-637

Query Match          24.8%; Score 422; DB 4; Length 145;
Best Local Similarity 100.0%; Pred. No. 7.6e-31;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 210 GLVGGRRYERPLVNGKVGWYTGWRADRPFAIDMAGFAVSLQVILSNPKAVFKRRGSQP 269
DB 23 GLVGGRRYERPLVNGKVGWYTGWRADRPFAIDMAGFAVSLQVILSNPKAVFKRRGSQP 82
QY 270 GMOESDFLQITVVEELEPKA 290
DB 83 GMOESDFLQITVVEELEPKA 103

RESULT 13
US-10-450-763-30691
; Sequence 30691, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 30691
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (583)...(632)
; OTHER INFORMATION: PHOSPHORYLASE KINASE ALPHA MUSCL domain identified by
; OTHER INFORMATION: EMATRIX, accession number PD01841F, p-value=7.048e-09,
; OTHER INFORMATION: 13.36
US-10-450-763-30691

Query Match          23.8%; Score 406; DB 5; Length 691;
Best Local Similarity 36.4%; Pred. No. 1.8e-28;
Matches 102; Conservative 28; Mismatches 56; Indels 94; Gaps 8;

QY 51 RLPLRRGGP-----AHGTQKRNS-----RPQP-----QPEPQ-LPTIYAIT 86
DB 224 RPPLRAGQCDCLPPLRLAAAEQLRQKDLRISQLQAE LRPPPPAPAPQPPPEALPTIYVVT 283
QY 87 PTYSRVPQKAE LTRLANTFRQVAQLHWILVEDAAARSELVSRFLARAGLPSTHLHVPTPR 146
DB 284 PTYARLVQKAE L-----
QY 147 RYKRPLGRATEORNAGLAWLR-----HOHQAQP-----GVLFFADDNTYSLELFOEM 198
DB 296 -----RNKALDWLRGRSANGVKGDPPSPGTGGVVYFADSGNTYSWELFEEM 342
QY 199 RTRTKYSVMPVGLVGRRYERPLVNGKVGWYTGWRADRPFAIDMAGFAVSLQVILSNP 258

```



GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2006, 10:55:30 ; Search time 18 Seconds  
(without alignments)  
830.554 Million cell updates/sec

Title: US-10-634-905-2  
Perfect score: 1705  
Sequence: 1 MKSALFTRFFILLPWILVI.....EKNLANEPKYHLDTVKIEV 323

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 235405 seqs, 46284737 residues

Total number of hits satisfying chosen parameters: 235405

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:

- 1: /SIDSS5/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /SIDSS5/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 3: /SIDSS5/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 4: /SIDSS5/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 5: /SIDSS5/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 6: /SIDSS5/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 7: /SIDSS5/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 8: /SIDSS5/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 9: /SIDSS5/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 10: /SIDSS5/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 11: /SIDSS5/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 12: /SIDSS5/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1705	100.0	323	9	US-10-330-773-566
2	1207	70.8	290	9	US-10-330-773-563
3	249	14.6	408	11	US-11-096-568A-18021
4	249	14.6	440	11	US-11-096-568A-18020
5	249	14.6	448	11	US-11-096-568A-18019
6	243.5	14.3	354	11	US-11-096-568A-2084
7	243.5	14.3	385	11	US-11-096-568A-2083
8	243.5	14.3	441	11	US-11-096-568A-2082
9	193.5	11.3	492	9	US-10-467-962B-51
10	117.5	6.9	96	9	US-10-475-075-809
11	99.5	5.8	813	8	US-10-511-814-16
12	99.5	5.8	813	9	US-10-661-966-19
13	95	5.6	424	11	US-11-096-568A-22088
14	94.5	5.5	296	11	US-11-096-568A-1548
15	94.5	5.5	392	11	US-11-096-568A-21010
16	94	5.5	527	11	US-11-096-568A-23555
17	93	5.5	783	8	US-10-509-131-25
18	93	5.5	783	11	US-11-186-284-59
19	92.5	5.4	721	11	US-11-186-284-67
20	92	5.4	480	9	US-10-330-773-786
21	90.5	5.3	278	11	US-11-096-568A-9264

22	90.5	5.3	278	11	US-11-096-568A-9266	Sequence 9266, Ap
23	90.5	5.3	319	11	US-11-096-568A-9263	Sequence 9263, Ap
24	90	5.3	254	11	US-11-096-568A-3889	Sequence 3889, Ap
25	90	5.3	293	11	US-11-096-568A-3888	Sequence 3888, Ap
26	89	5.2	461	11	US-11-096-568A-18013	Sequence 18013, A
27	89	5.2	462	11	US-11-096-568A-18012	Sequence 18012, A
28	89	5.2	463	11	US-11-096-568A-18011	Sequence 18011, A
29	89	5.2	614	11	US-11-015-546A-20	Sequence 20, Appl
30	88	5.2	221	11	US-11-096-568A-17196	Sequence 17196, A
31	88	5.2	233	11	US-11-096-568A-17195	Sequence 17195, A
32	88	5.2	1189	11	US-11-182-885-3	Sequence 3, Appli
33	87.5	5.1	250	11	US-11-096-568A-21169	Sequence 21169, A
34	87.5	5.1	409	11	US-11-096-568A-9548	Sequence 9548, Ap
35	87	5.1	416	11	US-11-096-568A-16843	Sequence 16843, A
36	87	5.1	528	11	US-11-087-099-7839	Sequence 7839, Ap
37	86.5	5.1	663	11	US-11-072-512-3771	Sequence 3771, Ap
38	86.5	5.1	1052	11	US-11-020-602-208	Sequence 208, App
39	86.5	5.1	2725	11	US-11-113-424-52	Sequence 52, Appl
40	86.5	5.1	2725	11	US-11-100-640-10	Sequence 16, Appl
41	86.5	5.1	2725	11	US-11-100-640-16	Sequence 10, Appl
42	86	5.0	534	11	US-11-096-568A-24457	Sequence 24457, A
43	86	5.0	548	11	US-11-096-568A-24456	Sequence 24456, A
44	85.5	5.0	390	11	US-11-096-568A-21913	Sequence 21913, A
45	85	5.0	281	11	US-11-096-568A-22579	Sequence 22579, A

ALIGNMENTS

RESULT 1

US-10-330-773-566

Sequence 566, Application US/10330773

Publication No. US20060040262A1

GENERAL INFORMATION:

APPLICANT: David W. Morris

APPLICANT: Marc Malandro

TITLE OF INVENTION: Novel Compositions and Methods in Cancer

FILE REFERENCE: 529452001300

CURRENT APPLICATION NUMBER: US/10/330,773

CURRENT FILING DATE: 2002-12-27

NUMBER OF SEQ ID NOS: 981

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 566

LENGTH: 323

TYPE: PRT

ORGANISM: Homo sapiens

US-10-330-773-566

Query Match 100.0%; Score 1705; DB 9; Length 323;

Best Local Similarity 100.0%; Pred. No. 1.9e-148;

Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MKSALFTRFFILLPWILVIIMLDVTRRPVPLTRPPSPYAVGGRGARLPURGGPA	60
Db	1	MKSALFTRFFILLPWILVIIMLDVTRRPVPLTRPPSPYAVGGRGARLPURGGPA	60
Qy	61	HGTQKNQSRPQPPQPPQPTTVAITPTYSRPQKAEITELANTERVOAHLWLVEDAA	120
Db	61	HGTQKNQSRPQPPQPPQPTTVAITPTYSRPQKAEITELANTERVOAHLWLVEDAA	120
Qy	121	ARSELVSRFLARAGLPSLTHVTPRRYKRPGLPRATEQRNAGLWLRQHQHQAQPGV	180
Db	121	ARSELVSRFLARAGLPSLTHVTPRRYKRPGLPRATEQRNAGLWLRQHQHQAQPGV	180
Qy	181	LPFADDNTYSLFQEMRTTRKVSVMVGLVGGRYERPLVNGKVGVYTGWRDRPF	240
Db	181	LPFADDNTYSLFQEMRTTRKVSVMVGLVGGRYERPLVNGKVGVYTGWRDRPF	240
Qy	241	AIDMAGFVSLQVILSNPKAVFKRGSCQMQSDFLKQITTVVELEPKANCTKVLVWH	300
Db	241	AIDMAGFVSLQVILSNPKAVFKRGSCQMQSDFLKQITTVVELEPKANCTKVLVWH	300
Qy	301	TRTEKVNLANEPKYHLDTVKIEV	323





```

; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 2082
; LENGTH: 441
; TYPE: PROT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(441)
; OTHER INFORMATION: Ceres Seq. ID no. 15180703
US-11-096-568A-2082

```

Query Match	14.3%	Score	243.5	DB	11	Length	441		
Best Local Similarity	28.9%	Pred. No.	2e-14						
Matches	76	Conservative	52	Mismatches	94	Indels	41	Gaps	12
Qy	76	EPQLPT---	IVAIPTYSRPVQKAELTRLANTFRQV-	AQLHWILVDEDAARSELVSRFLA	131				
Db	178	ESDIPVKKLLIIIVTITSVRPQAYLNLRAHTLKAVQAPLLWLIVE-WPEGSYETAELR	236						
Qy	132	RAGLPSTHL-----HVPTPRRYKRCLPRATEORNAGLAWLQRQHOBAPQCVLFPADDD	187						
Db	237	SSGVMYHMLCMRKNVTSVRKI-----AVCQRNAIYHVKRHH-----LDGIMHFADEE	284						
Qy	188	NTYSLELFEQEMRTTRKYSVMPVLVGRRY-----ERPLVENGKYGVTGTWRAD-----R	238						
Db	285	RSYSADVFEEQKIRRGSPMPVTIHICTKYRAVLGEPICKANRVMRWHTTQTQAKKSSTR	344						
Qy	239	PFAIDMAGFAVSLQVILSNPK-----AVFKRRGSGQPQWESDFLKQITTYE-ELEPK	289						
Db	345	RFPIGFSAFAFN-STLLWDQQRNRPMDMSVIVHTGGRGGLQESRFIEKLVKSRQRIEGL	403						
Qy	290	ANNCTKVLVWHHTREKYNLANEP	312						
Db	404	PDNCNRVMVWNFNLEPPOL-NYP	425						

```

RESULT 9
US-10-467-962B-51
US-10-467-962B-51
; Sequence 51, Application US/10467962B
; Publication No. US20050246784A1
; GENERAL INFORMATION:
; APPLICANT: Plesch, Gunnar
; APPLICANT: Blau, Astrid
; APPLICANT: Daeschner, Klaus
; APPLICANT: Klein, Mathieu
; TITLE OF INVENTION: Identification of Herbicidally Active Substances
; FILE REFERENCE: 2000 857
; CURRENT APPLICATION NUMBER: US/10/467,962B
; CURRENT FILING DATE: 2003-08-14
; PRIOR APPLICATION NUMBER: PCT/EP02/01466
; PRIOR FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn Vers. 2.0
; SEQ ID NO 51
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-467-962B-51

```

Qy	78	QL-----PTTVAITPTSRPVQKAEI	TRFLANTFRQVA-QLHWILVEDA	119
		:         :       :       :	:         :       :       :	
Db	133	QLLVRVQKEQKSGMYGRSPRTV	VVPTTYRTTQALHLTGVMHSLMLVPD	192
		:         :       :       :	:         :       :       :	
Qy	120	AARSELVSRFLARAGLPSTH	LVPTPRRYKRPGLPRATEORNAGLAWLR	176
		:         :       :       :	:         :       :       :	
Db	193	GITNETAS-FIAKSGLUKTIHL	GFQDK-----MPNTWEDRHKLKTYKMR	244
		:         :       :       :	:         :       :       :	
Qy	177	QPGVLPFADDNDNTYSLELFO	EMTTTRKVSVMVPGVLVG-----GRYERP	220
		:         :       :       :	:         :       :       :	
Db	245	LDGIVMFADDNSNMHSMELF	DEIQTVKWFAGALSVGILAHSADELS	304
		:         :       :       :	:         :       :       :	
Qy	221	LV-----ENGKVGWGY-----	TGWRADR-----PFAIDMAGFAVSLQWILSN	257
		:         :       :       :	:         :       :       :	
Db	305	SMPTQGPSCNSSEKLGVGHIF	NTQPYAKKTAVYIDEKAPVMPKMEWSGFV	357
		:         :       :       :	:         :       :       :	
Qy	258	PKAVFKRR-CSQFCMQESDPL	KQITVVEE-----LEPKANCTKVLVW	299
		:         :       :       :	:         :       :       :	
Db	358	SRLWKESSLDKPA-----	WVKDLSLLDDGYAIESPLSLVKDPSMVEPIG	412
		:         :       :       :	:         :       :       :	
Qy	300	HTRTE 304		
		:         :       :       :	:         :       :       :	
Db	413	WLRYE 417		
		:         :       :       :	:         :       :       :	
RESULT 10				
US-10-475-075-809				
; Sequence 809, Application US/10475075				
; Publication No. US20060053498A1				
; GENERAL INFORMATION:				
; APPLICANT: Bejanin, Stephane				
; APPLICANT: Tanaka, Hiroaki				
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste				
; APPLICANT: Jobert, Severin				
; APPLICANT: Giordano, Jean-Yves				
; TITLE OF INVENTION: Full-length human cDNAs encoding potentially secreted proteins				
; FILE REFERENCE: G-08IUS03PCT				
; CURRENT APPLICATION NUMBER: US/10/475,075				
; CURRENT FILING DATE: 2003-10-17				
; PRIOR APPLICATION NUMBER: PCT/IB01/00914				
; PRIOR FILING DATE: 2001-04-18				
; NUMBER OF SEQ ID NOS: 918				
; SOFTWARE: Patent.pm				
; SEQ ID NO 809				
; LENGTH: 96				
; TYPE: PRT				
; ORGANISM: Homo sapiens				
; FEATURE:				
; NAME/KEY: SIGNAL				
; LOCATION: -33..-1				
US-10-475-075-809				
Query Match 6.9%; Score 117.5; DB 9; Length 96;				
Best Local Similarity 31.4%; Pred. No. 0.00095;				
Matches 37; Conservative 13; Mismatches 25; Indels 43; Gaps 6;				
Qy	1	MKSALFTRFPILLPWILVIIM	LDVDTRRP---VPPLTPRPYSPYAVGRGA	50
		:         :       :       :	:         :       :       :	
Db	5	LKNVFLAYFLVSTAGLLYALV	QLG-----QPDCCLPPL-----RAAAEQLRQKD	48
		:         :       :       :	:         :       :       :	
Qy	51	-----RLPLRRGGPAHGTQ	KRNSRQPQPEQ-LFTIYAITPTYSRPVQKAE	101
		:         :       :       :	:         :       :       :	
Db	49	LRIISOIAEILRRPPPA-----	PAQPEPEALPTIYVVTPTVARLVYQKAE	96
		:         :       :       :	:         :       :       :	





GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2006, 10:55:51 ; Search time 57 Seconds  
(without alignments)  
2489.813 Million cell updates/sec

Title: US-10-634-905-2

Perfect score: 323

Sequence: 1 MKSALETRFPILLPWILIVI.....EKNLANEPKHYLDTVKIEV 323

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2443163 seqs, 439378781 residues

Word size : 1

Total number of hits satisfying chosen parameters: 2442881

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

A\_Geneseq 21:\*

1: Geneseq1980s:\*

2: Geneseq1990s:\*

3: Geneseq2000s:\*

4: Geneseq2001s:\*

5: Geneseq2002s:\*

6: Geneseq2003as:\*

7: Geneseq2003bs:\*

8: Geneseq2004s:\*

9: Geneseq2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	323	100.0	323	5	ADI03930 Human enz
2	323	100.0	323	8	Adq97589 Human can
3	197	61.0	197	5	ADI03933 HumanUDP-
4	81	25.1	145	4	Aam99900 Human exc
5	81	25.1	145	4	Aam43699 Human bla
6	81	25.1	145	8	Adf71658 Human bla
7	54	16.7	292	8	Adq97586 Mouse can
8	51	15.8	59	3	Aab45114 Human sec
9	47	14.6	324	5	ADI03932 Rat galac
10	28	8.7	29	3	AAB45116 Human sec
11	15	4.6	334	6	ABO07264 Human p53
12	15	4.6	347	2	AAy08214 Rat glucu
13	15	4.6	347	2	AAW90057 Rat GlCAT
14	15	4.6	347	5	ABB97603 Novel hum
15	13	4.0	40	8	ABO59511 Human gen
16	12	3.7	167	4	ABg13553 Novel hum
17	11	3.4	22	2	AAW90060 Rat GlCAT
18	11	3.4	241	8	AAW93879 Human pol
19	11	3.4	241	8	ADL31964 Human pro
20	11	3.4	335	6	ABO07265 Human p53
21	11	3.4	335	8	ADQ31194 Human GT
22	11	3.4	335	8	ADQ21418 Human sof
23	11	3.4	335	9	ADY70386 Galactosy
24	11	3.4	691	4	ABG00332 Novel hum

25	10	3.1	59	3	AAB45113	Aab45113 Human sec
26	10	3.1	201	8	ADO57267	Ado57267 Kidney de
27	10	3.1	306	8	ADI00832	Adi00832 Fruit fly
28	9	2.8	29	3	AAB45115	Aab45115 Human sec
29	8	2.5	25	2	AAW90059	Aaw90059 Rat GlCAT
30	8	2.5	107	6	ADA34275	Ada34275 Acinetoba
31	8	2.5	144	6	ABP97828	Abp97828 Amino aci
32	8	2.5	256	7	ABO76924	Abp76924 Pseudomon
33	8	2.5	272	1	AAp93560	Aap93560 Plasmodiu
34	8	2.5	316	8	ADI00836	Adi00836 Fruit fly
35	8	2.5	405	7	ABO68141	Abp68141 Pseudomon
36	8	2.5	410	7	ABO81824	Abp81824 Pseudomon
37	8	2.5	479	4	ABB62172	Abb62172 Drosophil
38	8	2.5	487	8	ADP08279	Adp08279 Neisseria
39	8	2.5	498	6	ABP79238	Abp79238 N. gonorr
40	8	2.5	530	7	ABO74593	Abp74593 Pseudomon
41	8	2.5	696	6	ABU08785	Abu08785 Transposa
42	8	2.5	912	8	ADJ76291	Adj76291 Marker ge
43	7	2.2	9	4	AAm53158	Aam53158 Human con
44	7	2.2	12	4	AAm00395	Aam00395 Human pro
45	7	2.2	19	4	AAm19546	Aam19546 Peptide #

## ALIGNMENTS

RESULT 1					
ADI03930					
ID	ADI03930 standard; protein; 323 AA.				
XX					
AC	ADI03930;				
XX					
DT	22-APR-2004 (first entry)				
XX					
DE	Human enzyme protein.				
XX					
KW	Enzyme protein; drug screening; disease diagnosis; human; gene therapy;				
KW	Chromosome 6; enzyme; glucuronyltransferase.				
XX					
OS	Homo sapiens.				
XX					
PN	WO200268657-A2.				
XX					
PD	06-SEP-2002.				
XX					
PF	08-FEB-2002; 2002WO-US003623.				
XX					
PR	26-FEB-2001; 2001US-0270871P.				
PR	26-MAR-2001; 2001US-00816095.				
XX	(PEKE ) PE CORP NY.				
PA	Gan W, Yan C, Merkulov GV, Di Francesco V, Beasley EM;				
XX					
PI	WPI; 2002-713380/77.				
XX					
DR	N-PSDB; ADI03929, ADI03931.				
DR					
XX					
PT	New human enzyme proteins, useful for treating or diagnosing disorders				
PT	associated with abnormal expression of the protein, in drug screening				
XX	assays and pharmacogenomic analysis.				
XX					
PS	Claim 1; SEQ ID NO 2; 127pp; English.				
XX					
CC	The invention relates to a novel isolated enzyme protein and encoding				
CC	polynucleotides. The protein shows a high degree of similarity to a				
CC	glucuronyltransferase cloned from a rabbit brain cDNA library. The				
CC	peptides and nucleic acid molecules are useful as models for the				
CC	development of human therapeutic targets, aid in the identification of				
CC	therapeutic proteins, and serve as targets for the development of human				
CC	therapeutic agents. The peptide may be used in drug screening assays, in				
CC	assays to determine the biological activity of the protein, to raise				
CC	antibodies or to elicit another immune response, as a reagent in assays				
CC	designed to quantitatively determine levels of the protein in biological				

CC fluids, or as markers for tissues in which the corresponding protein is  
CC preferentially expressed. The human enzyme protein is also useful for  
CC diagnosing a disease, predisposition to a disease, or treating a disorder  
CC characterized by an absence of, inappropriate or unwanted expression of  
CC the protein. The antibodies are useful in pharmacogenomic analysis, for  
CC inhibiting protein function, or for tissue typing. The nucleic acid  
CC molecules are useful as probes, primers, chemical intermediates, or in  
CC biological assays. The present sequence represents a human enzyme  
CC protein.  
XX  
SQ Sequence 323 AA;

Query Match 100.0%; Score 323; DB 5; Length 323;  
Best Local Similarity 100.0%; Pred. No. 2.7e-315;  
Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKSALFTRFFILLPWILVILIMLDVTRRPVPLTPRPYSPYAVGCGARLPLRGGPA 60  
DB 1 MKSALFTRFFILLPWILVILIMLDVTRRPVPLTPRPYSPYAVGCGARLPLRGGPA 60  
QY 61 HGTQKNQSRPQPEPQLPTIYAITPTYSRPVQKAELTRLANTFRQAQLHWILVEDAA 120  
DB 61 HGTQKNQSRPQPEPQLPTIYAITPTYSRPVQKAELTRLANTFRQAQLHWILVEDAA 120  
QY 121 ARSELVSRFLARAGLPSHLHVTPRRYKRPGLPRATEQORNAGLAWLRQHQRAPQGV 180  
DB 121 ARSELVSRFLARAGLPSHLHVTPRRYKRPGLPRATEQORNAGLAWLRQHQRAPQGV 180  
QY 181 LFFADDDNTYSLELFQEMRTTRKVSVPVGLVGGRRYERPLVGVGMYTGWADRPF 240  
DB 181 LFFADDDNTYSLELFQEMRTTRKVSVPVGLVGGRRYERPLVGVGMYTGWADRPF 240  
QY 241 AIMAGFAVSLQVILSNPKAVFKRGSQPMQSDFLKQITTVVELEPKANNCTKVLVWH 300  
DB 241 AIMAGFAVSLQVILSNPKAVFKRGSQPMQSDFLKQITTVVELEPKANNCTKVLVWH 300  
QY 301 TRTEKVNLANEPKYHLDTVKIEV 323  
DB 301 TRTEKVNLANEPKYHLDTVKIEV 323

RESULT 2  
ADQ97589  
ID ADQ97589 standard; protein; 323 AA.  
XX  
AC ADQ97589;  
XX  
DT 07-OCT-2004 (first entry)  
XX  
DE Human cancer associated sequence HP10-009, SEQ ID 566.  
XX  
KW Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human.  
XX  
OS Homo sapiens.  
XX  
FN WO2004060304-A2.  
XX  
PD 22-JUL-2004.  
XX  
PF 22-DEC-2003; 2003WO-US041389.  
XX  
PR 27-DEC-2002; 2002US-00330773.  
XX  
PA (SAGR-) SAGRES DISCOVERY INC.  
XX  
PI Morris DW, Malandro MS;  
XX  
DR WPI; 2004-543781/52.  
XX  
PT New isolated cancer associated nucleic acids comprising at least 10  
PT contiguous nucleotides, useful for diagnosing, preventing and/or treating  
PT cancers such as leukemia and lymphoma.  
XX

PS Claim 1; SEQ ID NO 566; 199pp; English.  
XX  
CC The present invention relates to cancer associated sequences (ADQ97025-  
CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or  
CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence  
CC data for this patent did not form part of the printed specification, but  
CC was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 323 AA;

Query Match 100.0%; Score 323; DB 8; Length 323;  
Best Local Similarity 100.0%; Pred. No. 2.7e-315;  
Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKSALFTRFFILLPWILVILIMLDVTRRPVPLTPRPYSPYAVGCGARLPLRGGPA 60  
DB 1 MKSALFTRFFILLPWILVILIMLDVTRRPVPLTPRPYSPYAVGCGARLPLRGGPA 60  
QY 61 HGTQKNQSRPQPEPQLPTIYAITPTYSRPVQKAELTRLANTFRQAQLHWILVEDAA 120  
DB 61 HGTQKNQSRPQPEPQLPTIYAITPTYSRPVQKAELTRLANTFRQAQLHWILVEDAA 120  
QY 121 ARSELVSRFLARAGLPSHLHVTPRRYKRPGLPRATEQORNAGLAWLRQHQRAPQGV 180  
DB 121 ARSELVSRFLARAGLPSHLHVTPRRYKRPGLPRATEQORNAGLAWLRQHQRAPQGV 180  
QY 181 LFFADDDNTYSLELFQEMRTTRKVSVPVGLVGGRRYERPLVGVGMYTGWADRPF 240  
DB 181 LFFADDDNTYSLELFQEMRTTRKVSVPVGLVGGRRYERPLVGVGMYTGWADRPF 240  
QY 241 AIMAGFAVSLQVILSNPKAVFKRGSQPMQSDFLKQITTVVELEPKANNCTKVLVWH 300  
DB 241 AIMAGFAVSLQVILSNPKAVFKRGSQPMQSDFLKQITTVVELEPKANNCTKVLVWH 300  
QY 301 TRTEKVNLANEPKYHLDTVKIEV 323  
DB 301 TRTEKVNLANEPKYHLDTVKIEV 323

RESULT 3  
ADI03933  
ID ADI03933 standard; protein; 197 AA.  
XX  
AC ADI03933;  
XX  
DT 22-APR-2004 (first entry)  
XX  
DE HumanUDP-glucuronyltransferase-S fragment.  
XX  
KW Enzyme protein; drug screening; disease diagnosis; rat; gene therapy;  
KW chromosome 6; enzyme; glucuronyltransferase; UDP-glucuronyltransferase-S.  
XX  
OS Homo sapiens.  
XX  
FN WO200268657-A2.  
XX  
PD 06-SEP-2002.  
XX  
PF 08-FEB-2002; 2002WO-US003623.  
XX  
PR 26-FEB-2001; 2001US-0270871P.  
PR 26-MAR-2001; 2001US-00816095.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Gan W, Yan C, Merkulov GV, Di Francesco V, Beasley EM;  
XX  
DR WPI; 2002-713380/77.  
XX  
PT New human enzyme proteins, useful for treating or diagnosing disorders  
PT associated with abnormal expression of the protein, in drug screening  
PT assays and pharmacogenomic analysis.  
XX

```
XX Disclosure; SEQ ID NO 5; 127pp; English.
XX The invention relates to a novel isolated enzyme protein and encoding
XX polynucleotides. The protein shows a high degree of similarity to a
XX glucuronyltransferase cloned from a rabbit brain cDNA library. The
XX peptides and nucleic acid molecules are useful as models for the
XX development of human therapeutic targets, aid in the identification of
XX therapeutic proteins, and serve as targets for the development of human
XX therapeutic agents. The peptide may be used in drug screening assays, in
XX assays to determine the biological activity of the protein, to raise
XX antibodies or to elicit another immune response, as a reagent in assays
XX designed to quantitatively determine levels of the protein in biological
XX fluids, or as markers for tissues in which the corresponding protein is
XX preferentially expressed. The human enzyme protein is also useful for
XX diagnosing a disease, predisposition to a disease, or treating a disorder
XX characterized by an absence of, inappropriate or unwanted expression of
XX the protein. The antibodies are useful in pharmacogenomic analysis, for
XX inhibiting protein function, or for tissue typing. The nucleic acid
XX molecules are useful as probes, primers, chemical intermediates, or in
XX biological assays. The present sequence represents a human UDP-
XX glucuronyltransferase-S fragment used in alignment studies with the novel
XX human enzyme protein.
XX Sequence 197 AA;
SQ
Query Match 61.0%; Score 197; DB 5; Length 197;
Best Local Similarity 100.0%; Pred. No. 5.6e-189; Indels 0; Gaps 0;
Matches 197; Conservative 0; Mismatches 0;
QY 1 MKSALFTRFFILLPWILVIMLDVTRRPVPLTPRPYFSPYAVGGGARLPRLRGGA 60
DB 1 MKSALFTRFFILLPWILVIMLDVTRRPVPLTPRPYFSPYAVGGGARLPRLRGGA 60
QY 61 HGTOKNQSRPQPOPEQLPTIYAITYSRPVQKAELTRANTFROAQLHWILVEDAA 120
DB 61 HGTOKNQSRPQPOPEQLPTIYAITYSRPVQKAELTRANTFROAQLHWILVEDAA 120
QY 121 ARSELVSRFLARAGLPSTHLVPTPRYKRPGLPRATEQRNAGLAWLRQHRHQRAQPGV 180
DB 121 ARSELVSRFLARAGLPSTHLVPTPRYKRPGLPRATEQRNAGLAWLRQHRHQRAQPGV 180
QY 181 LFFADDDNTYSLELFOE 197
DB 181 LFFADDDNTYSLELFOE 197
RESULT 4
AAM99900
ID AAM99900 standard; protein; 145 AA.
XX AC AAM99900;
XX 07-JAN-2002 (first entry)
XX Human excretory related polypeptide SEQ ID NO 637.
XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
XX immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
XX antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
XX antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
XX antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
XX antiparasitic; cardiant; immune disorder; cardiovascular disorder;
XX neurological disease; infection; nephrotropic; gene therapy; vaccine;
XX excretory system.
XX Homo sapiens.
XX WO200155313-A2.
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US001323.
XX
```

```
XX 31-JAN-2000; 2000US-0179065P. XX
PR 04-FEB-2000; 2000US-0180628P. PR
PR 24-FEB-2000; 2000US-0184664P. PR
PR 02-MAR-2000; 2000US-0186350P. PR
PR 16-MAR-2000; 2000US-0189874P. PR
PR 17-MAR-2000; 2000US-0190076P. PR
PR 18-APR-2000; 2000US-0198123P. PR
PR 19-MAY-2000; 2000US-0205515P. PR
PR 07-JUN-2000; 2000US-0209467P. PR
PR 28-JUN-2000; 2000US-0214886P. PR
PR 30-JUN-2000; 2000US-0215135P. PR
PR 07-JUL-2000; 2000US-0216647P. PR
PR 11-JUL-2000; 2000US-0216880P. PR
PR 07-JUL-2000; 2000US-0217487P. PR
PR 11-JUL-2000; 2000US-0217496P. PR
PR 14-JUL-2000; 2000US-0218290P. PR
PR 26-JUL-2000; 2000US-0220963P. PR
PR 26-JUL-2000; 2000US-0220964P. PR
PR 14-AUG-2000; 2000US-0224518P. PR
PR 14-AUG-2000; 2000US-0224519P. PR
PR 14-AUG-2000; 2000US-0225213P. PR
PR 14-AUG-2000; 2000US-0225214P. PR
PR 14-AUG-2000; 2000US-0225266P. PR
PR 14-AUG-2000; 2000US-0225267P. PR
PR 14-AUG-2000; 2000US-0225268P. PR
PR 14-AUG-2000; 2000US-0225270P. PR
PR 14-AUG-2000; 2000US-0225447P. PR
PR 14-AUG-2000; 2000US-0225757P. PR
PR 14-AUG-2000; 2000US-0225758P. PR
PR 14-AUG-2000; 2000US-0225759P. PR
PR 18-AUG-2000; 2000US-022679P. PR
PR 22-AUG-2000; 2000US-0226681P. PR
PR 22-AUG-2000; 2000US-0226688P. PR
PR 22-AUG-2000; 2000US-0227182P. PR
PR 23-AUG-2000; 2000US-0227009P. PR
PR 30-AUG-2000; 2000US-0228924P. PR
PR 01-SEP-2000; 2000US-0229287P. PR
PR 01-SEP-2000; 2000US-0229343P. PR
PR 01-SEP-2000; 2000US-0229344P. PR
PR 01-SEP-2000; 2000US-0229345P. PR
PR 05-SEP-2000; 2000US-0229509P. PR
PR 05-SEP-2000; 2000US-0229513P. PR
PR 06-SEP-2000; 2000US-0230437P. PR
PR 06-SEP-2000; 2000US-0230438P. PR
PR 08-SEP-2000; 2000US-0231242P. PR
PR 08-SEP-2000; 2000US-0231243P. PR
PR 08-SEP-2000; 2000US-0231244P. PR
PR 08-SEP-2000; 2000US-0231413P. PR
PR 08-SEP-2000; 2000US-0231414P. PR
PR 08-SEP-2000; 2000US-0232080P. PR
PR 08-SEP-2000; 2000US-0232081P. PR
PR 12-SEP-2000; 2000US-0231968P. PR
PR 14-SEP-2000; 2000US-0232397P. PR
PR 14-SEP-2000; 2000US-0232398P. PR
PR 14-SEP-2000; 2000US-0232399P. PR
PR 14-SEP-2000; 2000US-0232399P. PR
PR 14-SEP-2000; 2000US-0232400P. PR
PR 14-SEP-2000; 2000US-0232401P. PR
PR 14-SEP-2000; 2000US-0233063P. PR
PR 14-SEP-2000; 2000US-0233064P. PR
PR 14-SEP-2000; 2000US-0233065P. PR
PR 21-SEP-2000; 2000US-0234223P. PR
PR 21-SEP-2000; 2000US-0234274P. PR
PR 25-SEP-2000; 2000US-0234997P. PR
PR 25-SEP-2000; 2000US-0234998P. PR
PR 26-SEP-2000; 2000US-0235484P. PR
PR 27-SEP-2000; 2000US-0235834P. PR
PR 27-SEP-2000; 2000US-0235836P. PR
PR 29-SEP-2000; 2000US-0236327P. PR
PR 29-SEP-2000; 2000US-0236367P. PR
PR 29-SEP-2000; 2000US-0236368P. PR
PR 29-SEP-2000; 2000US-0236369P. PR
PR 29-SEP-2000; 2000US-0236370P. PR
```

PR	02-OCT-2000;	2000US-0236802P.	PT	in preventing, treating or ameliorating a medical condition.
PR	02-OCT-2000;	2000US-0237037P.	XX	
PR	02-OCT-2000;	2000US-0237038P.	PS	Claim 11; SEQ ID NO 637; 574pp + Sequence Listing; English.
PR	02-OCT-2000;	2000US-0237039P.	XX	
PR	12-OCT-2000;	2000US-0237040P.	CC	The invention relates to novel excretory system related human
PR	13-OCT-2000;	2000US-0239935P.	CC	polynucleotides (AAI98567-AAI99503) and the encoded proteins (AAM99594-
PR	13-OCT-2000;	2000US-0239937P.	CC	AAM99913) useful for preventing, treating or ameliorating medical
PR	20-OCT-2000;	2000US-0240960P.	CC	conditions e.g. by protein or gene therapy, especially disorders related
PR	20-OCT-2000;	2000US-0241221P.	CC	to the excretory system. The genes are isolated from a range of human
PR	20-OCT-2000;	2000US-0241785P.	CC	tissues disclosed in the specification. The nucleic acids, proteins, and
PR	20-OCT-2000;	2000US-0241787P.	CC	antibodies and (ant)agonists are useful in the diagnosis, treatment and
PR	20-OCT-2000;	2000US-0241808P.	CC	prevention of: (a) cancer, e.g. breast and ovarian cancer and other
PR	20-OCT-2000;	2000US-0241809P.	CC	cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal
PR	20-OCT-2000;	2000US-0241826P.	CC	tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's
PR	01-NOV-2000;	2000US-0244617P.	CC	disease, allergies, autoimmune haemolytic anaemia, autoimmune
PR	08-NOV-2000;	2000US-0246474P.	CC	thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis,
PR	08-NOV-2000;	2000US-0246475P.	CC	rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders
PR	08-NOV-2000;	2000US-0246476P.	CC	such as myocardial ischaemias; (d) wound healing; (e) neurological
PR	08-NOV-2000;	2000US-0246477P.	CC	diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases
PR	08-NOV-2000;	2000US-0246478P.	CC	such as viral, bacterial, fungal and parasitic infections. Note: The
PR	08-NOV-2000;	2000US-0246523P.	CC	sequence data for this patent did not form part of the printed
PR	08-NOV-2000;	2000US-0246524P.	CC	specification, but was obtained in electronic format directly from WIPO
PR	08-NOV-2000;	2000US-0246525P.	CC	at ftp.wipo.int/pub/published_pct_sequences
PR	08-NOV-2000;	2000US-0246526P.	XX	
PR	08-NOV-2000;	2000US-0246527P.	SQ	Sequence 145 AA;
PR	08-NOV-2000;	2000US-0246528P.		
PR	08-NOV-2000;	2000US-0246532P.		
PR	08-NOV-2000;	2000US-0246609P.		
PR	08-NOV-2000;	2000US-0246610P.		
PR	08-NOV-2000;	2000US-0246611P.		
PR	08-NOV-2000;	2000US-0246613P.		
PR	17-NOV-2000;	2000US-0249207P.	QY	210 GLVGGRRYERPLVNGKVVGYTGTGRADRPFAIDMAGFAVSLQVILSNPKAVFKRGSGP 269
PR	17-NOV-2000;	2000US-0249208P.	DB	23 GLVGGRRYERPLVNGKVVGYTGTGRADRPFAIDMAGFAVSLQVILSNPKAVFKRGSGP 82
PR	17-NOV-2000;	2000US-0249209P.		
PR	17-NOV-2000;	2000US-0249210P.	QY	270 GMOESDFLQIITVVELEPKA 290
PR	17-NOV-2000;	2000US-0249211P.	DB	83 GMOESDFLQIITVVELEPKA 103
PR	17-NOV-2000;	2000US-0249212P.		
PR	17-NOV-2000;	2000US-0249213P.		
PR	17-NOV-2000;	2000US-0249214P.		
PR	17-NOV-2000;	2000US-0249215P.		
PR	17-NOV-2000;	2000US-0249216P.		
PR	17-NOV-2000;	2000US-0249217P.		
PR	17-NOV-2000;	2000US-0249218P.		
PR	17-NOV-2000;	2000US-0249244P.		
PR	17-NOV-2000;	2000US-0249245P.		
PR	17-NOV-2000;	2000US-0249264P.		
PR	17-NOV-2000;	2000US-0249265P.		
PR	17-NOV-2000;	2000US-0249297P.		
PR	17-NOV-2000;	2000US-0249299P.		
PR	01-DEC-2000;	2000US-0249300P.		
PR	01-DEC-2000;	2000US-0250160P.		
PR	01-DEC-2000;	2000US-0250391P.		
PR	05-DEC-2000;	2000US-0251030P.		
PR	05-DEC-2000;	2000US-0251988P.		
PR	05-DEC-2000;	2000US-0256719P.		
PR	06-DEC-2000;	2000US-0251479P.		
PR	08-DEC-2000;	2000US-0251856P.		
PR	08-DEC-2000;	2000US-0251868P.		
PR	08-DEC-2000;	2000US-0251869P.		
PR	08-DEC-2000;	2000US-0251989P.		
PR	08-DEC-2000;	2000US-0251990P.		
PR	11-DEC-2000;	2000US-0254097P.		
PR	05-JAN-2001;	2001US-0259678P.		
PA	(HUMA-)	HUMAN GENOME SCI INC.		
XX				
XX	Rosen CA, Barash SC, Ruben SM;			
XX	WPI; 2001-465569/50.			
DR	N-PSDB; AAI98873.			
DR				
XX	Isolated nucleic acid molecule encoding excretory system antigen is used			
PT				
Query Match	25.1%;	Score 81;	DB 4;	Length 145;
Best Local Similarity	100.0%;	Pred. No. 1.3e-72;		
Matches	81;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	210	GLVGGRRYERPLVNGKVVGYTGTGRADRPFAIDMAGFAVSLQVILSNPKAVFKRGSGP	269	
DB	23	GLVGGRRYERPLVNGKVVGYTGTGRADRPFAIDMAGFAVSLQVILSNPKAVFKRGSGP	82	
QY	270	GMOESDFLQIITVVELEPKA	290	
DB	83	GMOESDFLQIITVVELEPKA	103	
RESULT 5				
AAM43699				
ID	AAM43699	standard; protein; 145 AA.		
XX	AAM43699;			
AC	AAM43699;			
XX				
DT	24-OCT-2001	(first entry)		
XX				
DE	Human bladder antigen, SEQ ID NO: 93.			
XX				
KW	Human; bladder antigen; cytostatic; immunosuppressive; nootropic;			
KW	neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;			
KW	antiinflammatory; antitumor; antitubercular; anticonvulsant; antibacterial;			
KW	antifungal; antiparasitic; cardiant; gene therapy; cancer;			
KW	immune disorder; cardiovascular disorder; wound healing; infection;			
XX	neurological disease.			
OS	Homo sapiens.			
XX				
PN	WO200159064-A2.			
XX				
PD	16-AUG-2001.			
XX				
PF	17-JAN-2001; 2001WO-US001342.			
XX				
PR	31-JAN-2000; 2000US-0179065P.			
PR	04-FEB-2000; 2000US-0180628P.			
PR	24-FEB-2000; 2000US-0184664P.			
PR	02-MAR-2000; 2000US-0186350P.			
PR	16-MAR-2000; 2000US-0189874P.			
PR	17-MAR-2000; 2000US-0190076P.			
PR	18-APR-2000; 2000US-0198123P.			
PR	19-MAY-2000; 2000US-0205515P.			
PR	07-JUN-2000; 2000US-0209467P.			



CC urogenital system; immune disorders such as Addison's disease, allergies,  
CC auto-immune haemolytic anaemia, auto-immune thyroiditis, diabetes mellitus,  
CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; cardiovascular disorders such as myocardial ischaemias; wound  
CC healing; neurological diseases such as cerebral anoxia and epilepsy; and  
CC infectious diseases such as viral, bacterial, fungal and parasitic  
CC infections. Numerous examples of each type of disorder are given in the  
CC specification. The polypeptides can also be used as a food additive or  
CC preservative to increase or decrease storage capabilities. The  
CC polynucleotides are useful for chromosome identification. They are also  
CC useful as probes for diagnosing or treating a disorder related to the  
CC female reproductive system, particularly breast and/or ovary cancer. The  
CC present sequence is a novel bladder antigen provided in the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 145 AA;

Query Match 25.1%; Score 81; DB 4; Length 145;  
Best Local Similarity 100.0%; Pred. No. 1.3e-72;  
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 210 GLVGGRRYERPLVENGKVGWYTGWRADRPFAIDMAGFAVSLQVILSNPKAVFKRGSGP 269  
Db 23 GLVGGRRYERPLVENGKVGWYTGWRADRPFAIDMAGFAVSLQVILSNPKAVFKRGSGP 82  
QY 270 GMOESDFLKQITVVELEPKA 290  
Db 83 GMOESDFLKQITVVELEPKA 103

RESULT 6  
ADF71658  
ID ADF71658 standard; protein; 145 AA.

XX AC ADF71658;  
XX DT 11-MAR-2004 (first entry)  
XX DE Human bladder associated antigen #38.  
XX KW autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;  
KW breast neoplasms; liver neoplasms; cardiovascular disorder;  
KW cardiac arrest; cerebrovascular disorder; cerebral ischaemia;  
KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;  
KW ocular disorder; corneal infection; wound healing;  
KW epithelial cell proliferation; skin aging; sunburn;  
KW organ transplantation; cell culture; tissue regeneration; chemotaxis;  
KW food additive; preservative; Human; bladder associated antigen.

XX OS Homo sapiens.

XX PN US2003199008-A1.

XX PD 23-OCT-2003.

XX 22-FEB-2002; 2002US-00080254.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225271P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.



Best Local Similarity 100.0%; Pred. No. 2.9e-45;  
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 237 DRPFAIDMAGFAVSLQVILSNPKAVFKRGSGQGMQESDFLKQITTVBELEPKA 290  
|||||  
Db 206 DRPFAIDMAGFAVSLQVILSNPKAVFKRGSGQGMQESDFLKQITTVBELEPKA 259  
|||||

RESULT 8  
ID AAB45114  
AC AAB45114 standard; protein; 59 AA.

XX AAB45114;

XX DT 12-FEB-2001 (first entry)

XX DE Human secreted protein encoded by gene 46 homologue.

XX KW Secreted protein; human; immunosuppressive; antiarthritic; antirheumatic;  
XX KW antiproliferative; cycostatic; cardiant; vasotropic; cerebroprotective;  
XX KW neutropic; neuroprotective; antibacterial; virucide; fungicide;  
XX KW ophthalmological; vulnerary; gene therapy; treatment; autoimmune disease;  
XX KW hyperproliferative disorder; cardiovascular disorder; ocular disorder;  
XX KW cerebrovascular disorder; nervous system disorder; infection; skin aging;  
XX KW wound healing; epithelial cell proliferation; transplantation.

XX OS Homo sapiens.

XX PN WO200058357-A1.

XX PD 05-OCT-2000.

XX PF 23-MAR-2000; 2000WO-US0007723.

XX PR 26-MAR-1999; 99US-0126506P.

XX PR 07-JAN-2000; 2000US-0174852P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ruben SM, Komatsoulis G;

XX WPI; 2000-611704/58.

XX PT Nucleic acid molecules encoding human secreted proteins, used in  
preventing, treating or ameliorating a disorder, e.g. Alzheimer's and  
Parkinson's diseases and cancers.

XX PS Disclosure; Page 413; 418pp; English.

XX CC This invention describes novel isolated nucleic acid molecules (I)  
encoding a human secreted proteins (II) which have immunosuppressive,  
antiarthritic, antirheumatic, antiproliferative, cycostatic, cardiant,  
vasotropic, cerebroprotective, neutropic, neuroprotective, antibacterial,  
virucide, fungicide, ophthalmological and vulnerary activity and can be  
used for gene therapy. (I) and (II) are used to prevent, treat or  
ameliorate a medical condition in e.g. humans, mice, rabbits, goats,  
horses, cats, dogs, chickens or sheep. (I) and (II) are also used in  
diagnosing a pathological condition or susceptibility to a pathological  
condition. The antibodies to (I) can also be used in alleviating  
symptoms associated with the disorders and in diagnostic immunoassays  
e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA).  
Disorders which are diagnosed or treated include autoimmune diseases e.g.  
rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the  
breast or liver, cardiovascular disorders e.g. cardiac arrest,  
cerebrovascular disorders e.g. cerebral ischemia, angiogenesis, nervous  
system disorders e.g. Alzheimer's disease, infections caused by bacteria,  
viruses and fungi and ocular disorders e.g. corneal infection. The  
polypeptides can also be used to aid wound healing and epithelial cell  
proliferation, to prevent skin aging due to sunburn, to maintain organs  
before transplantation, for supporting cell culture of primary tissues,  
to regenerate tissues and in chemotaxis. The polypeptides can also be  
used as a food additive or preservative to increase or decrease storage  
capabilities

XX SQ Sequence 59 AA;

Query Match 15.8%; Score 51; DB 3; Length 59;  
Best Local Similarity 100.0%; Pred. No. 7.7e-43;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 246 GFAVSLQVILSNPKAVFKRGSGQGMQESDFLKQITTVBELEPKANNCTKV 296  
|||||  
Db 7 GFAVSLQVILSNPKAVFKRGSGQGMQESDFLKQITTVBELEPKANNCTKV 57  
|||||

RESULT 9

AD103932

ID AD103932 standard; protein; 324 AA.

XX AC AD103932;

XX DT 22-APR-2004 (first entry)

XX DE Rat galactosidase beta-1,3-glucuronyltransferase.

XX KW Enzyme protein; drug screening; disease diagnosis; rat; gene therapy;  
XX KW chromosome 6; enzyme; glucuronyltransferase;  
XX KW galactosidase beta-1,3-glucuronyltransferase.

XX OS Rattus norvegicus.

XX PN WO200268657-A2.

XX PD 06-SEP-2002.

XX PF 08-FEB-2002; 2002WO-US003623.

XX PR 26-FEB-2001; 2001US-0270871P.

XX PR 26-MAR-2001; 2001US-00816095.

XX PA (PEKE ) PE CORP NY.

XX PI Gan W, Yan C, Merkulov GV, Di Francesco V, Beasley EM;

XX WPI; 2002-713380/77.

XX REFSEQ; NP\_072131.1.

XX PT New human enzyme proteins, useful for treating or diagnosing disorders  
associated with abnormal expression of the protein, in drug screening  
assays and pharmacogenomic analysis.

XX PS Disclosure; SEQ ID NO 4; 127pp; English.

XX CC The invention relates to a novel isolated enzyme protein and encoding  
polynucleotides. The protein shows a high degree of similarity to a  
glucuronyltransferase cloned from a rabbit brain cDNA library. The  
peptides and nucleic acid molecules are useful as models for the  
development of human therapeutic targets, aid in the identification of  
therapeutic proteins, and serve as targets for the development of human  
therapeutic agents. The peptide may be used in drug screening assays, in  
assays to determine the biological activity of the protein, to raise  
antibodies or to elicit another immune response, as a reagent in assays  
designed to quantitatively determine levels of the protein in biological  
fluids, or as markers for tissues in which the corresponding protein is  
preferentially expressed. The human enzyme protein is also useful for  
diagnosing a disease, predisposition to a disease, or treating a disorder  
characterized by an absence of, inappropriate or unwanted expression of  
the protein. The antibodies are useful in pharmacogenomic analysis, for  
inhibiting protein function, or for tissue typing. The nucleic acid  
molecules are useful as probes, primers, chemical intermediates, or in  
biological assays. The present sequence represents a rat galactosidase  
beta-1,3-glucuronyltransferase used in alignment studies with the novel  
human enzyme protein.

XX SQ Sequence 324 AA;



CC Modulators identified by (M1) are useful in a variety of diagnostic and  
 CC therapeutic applications, where disease or disorder prognosis is related  
 CC to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell  
 CC proliferation disorders (e.g. cancer). Another two new methods (M2 and  
 CC M3) are useful for modulating the p53 pathway of a cell, thus restoring  
 CC the p53 function of the cell, so that the cell undergoes normal  
 CC proliferation or progression through the cell cycle. (M2) and (M3) are  
 CC also useful for treating defects in the p53 pathway such as angiogenic,  
 CC apoptotic or cell proliferation disorders. The present sequence  
 CC represents a human p53 pathway modifying protein  
 XX  
 SQ Sequence 334 AA;

Query Match 4.6%; Score 15; DB 6; Length 334;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 TPTYSRPVQKAELTR 100  
 |||||  
 Db 90 TPTYSRPVQKAELTR 104

RESULT 12  
 AAY08214  
 ID AAY08214 standard; protein; 347 AA.

XX  
 AC AAY08214;

XX  
 DT 09-JUL-1999 (first entry)

DE Rat glucuronyltransferase protein.

XX HNK-1; sulphotransferase; rat; enzyme; carbohydrate epitope; nCAM;  
 KW neural cell adhesion molecule; neural growth; remyelination; disease;  
 KW neural protection; treatment; neuron; glucuronyltransferase.

XX Rattus sp.

XX WO9911796-A1.

XX 11-MAR-1999.

XX 04-SEP-1998; 98WO-US018572.

XX 05-SEP-1997; 97US-00924812.

XX (ACOR-) ACORDA THERAPEUTICS.

XX Mantei N, Bakker H, Schachner M;

XX WPI; 1999-312405/26.

DR N-PSDB; AAX37750.

XX New enzyme useful for promoting neural growth, remyelination and neural  
 PT protection.

PS Disclosure; Page 80-82; 85pp; English.

XX This invention describes a novel isolated rat sulphotransferase enzyme  
 CC capable of generating the HNK-1 carbohydrate epitope on a neural cell  
 CC adhesion molecule (nCAM). HNK-1 sulphotransferase is used to promote  
 CC neural growth, remyelination, and/or neural protection in mammals. The  
 CC cDNA encoding the HNK-1 sulphotransferase enzyme is useful in the  
 CC treatment of diseased or damaged neurons. The HNK-1 sulphotransferase  
 CC cDNA and protein are also useful in screening for potential drugs  
 CC effective to modulate the sulphotransferase activity of target mammalian  
 CC cells. The product of the invention has neuroprotective activity. This  
 CC sequence represents a rat glucuronyltransferase protein used to describe  
 CC the method of the invention

XX Sequence 347 AA;

XX Query Match 4.6%; Score 15; DB 2; Length 347;

Best Local Similarity 100.0%; Pred. No. 4.8e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 TPTYSRPVQKAELTR 100  
 |||||  
 Db 103 TPTYSRPVQKAELTR 117

RESULT 13  
 AAW90057  
 ID AAW90057 standard; protein; 347 AA.

XX  
 AC AAW90057;

XX 26-FEB-1999 (first entry)

XX Rat GlcAT-P protein.

DE GlcAT-P; glucuronate transferase; rat; glucuronic acid; donor; receptor.

XX Rattus sp.

XX JP10313867-A.

XX 02-DEC-1998.

XX 16-MAY-1997; 97JP-00127065.

XX 16-MAY-1997; 97JP-00127065.

XX (SEK ) SEIKAGAKU KOGYO CO LTD.

XX WPI; 1999-074149/07.

DR N-PSDB; AAV73869.

XX DNA coding glucuronate transferase - comprises transferring glucuronic  
 PT acid from glucuronic acid donor to glucuronic acid receptor.

XX Claim 3; Page 13-14; 18pp; Japanese.

XX This sequence represents a novel rat glucuronate transferase (GlcAT-P)  
 CC which transfers glucuronic acid from a glucuronic acid donor to a  
 CC glucuronic acid receptor with an optimum reaction pH of 6.0 to 6.5. The  
 CC protein transfers specifically glucuronic acid to the N-  
 CC acetylglucosamine residue of asialo-orosomucoid and neuron adhesive  
 CC molecule, has a molecular weight of approximately 45,000 and maintains  
 CC its maintains its activity in the presence of 5 mM neolactotetraose-  
 CC phenylC14H29

XX Sequence 347 AA;

Query Match 4.6%; Score 15; DB 2; Length 347;

Best Local Similarity 100.0%; Pred. No. 4.8e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 TPTYSRPVQKAELTR 100  
 |||||  
 Db 103 TPTYSRPVQKAELTR 117

RESULT 14  
 ABB97603  
 ID ABB97603 standard; protein; 347 AA.

XX ABB97603;

XX 27-JUN-2002 (first entry)

XX Novel human protein SEQ ID NO: 871.

XX Human; antianaemic; vulnery; antiinflammatory; immunomodulator;  
 KW antinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;  
 KW neuroprotective; antiparkinsonian; protein therapy; EST;

KW expressed sequence tag.  
 XX Homo sapiens.  
 XX WO200222660-A2.  
 XX 21-MAR-2002.  
 XX 10-SEP-2001; 2001WO-US026015.  
 XX 11-SEP-2000; 2000US-00659671.  
 XX (HYSE-) HYSEQ INC.  
 XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
 XX Xue AJ, Yang Y, Wehrman T, Drmanac RT;  
 XX WPI; 2002-292408/33.  
 XX N-PSDB; ABN32789.  
 XX An isolated polynucleotide for treating diseases associated with its  
 XX encoded polypeptide such as cancer and multiple sclerosis.  
 XX Example 2; SEQ ID NO 871; 509pp; English.  
 XX The present invention provides the protein and coding sequences of 444  
 XX novel human proteins. These were isolated from expressed sequences tags  
 XX (ESTs). They can be used to stimulate cell growth, to regulate  
 XX haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth  
 XX e.g. in burn treatment, to regulate the immune system e.g. to treat  
 XX multiple sclerosis, to regulate activin or inhibit e.g. to treat  
 XX infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke  
 XX and cancer, to screen for drugs, to treat inflammatory conditions e.g.  
 XX rheumatoid arthritis, and to treat nervous system disorders e.g.  
 XX Parkinson's disease. The present sequence is a protein of the invention  
 XX SQ Sequence 347 AA;  
 Query Match 4.6%; Score 15; DB 5; Length 347;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 86 TPTYSRPVQKALTR 100  
 Db 103 TPTYSRPVQKALTR 117  
 RESULT 15  
 ABO59511  
 ID ABO59511 standard; protein; 40 AA.  
 XX ABO59511;  
 XX 29-JUL-2004 (first entry)  
 XX Human genome derived single exon protein #5745.  
 XX Human; gene expression; single exon probe; microarray;  
 XX alternative splicing event; genomic alteration.  
 XX Homo sapiens.  
 XX US2003194704-A1.  
 XX 16-OCT-2003.  
 XX 03-APR-2002; 2002US-00029386.  
 XX 03-APR-2002; 2002US-00029386.  
 XX (PENN/) PENN S G.  
 XX (RANK/) RANK D R.  
 XX (HANZ/) HANZEL D K.

XX Penn SG, Rank DR, Hanzel DK;  
 XX WPI; 2004-119264/12.  
 XX New human genome-derived single exon nucleic acid probes useful for human  
 XX gene expression analysis, for identifying or characterizing alternative  
 XX splicing events, for assessing genomic alterations or as tools for  
 XX surveying tissues.  
 XX Claim 45; SEQ ID NO 33145; 80pp; English.  
 XX The invention relates to a nucleic acid probe for measuring human gene  
 XX expression, comprising any of the 27,400 fully defined nucleotide  
 XX sequences in the specification, or their complements or fragments, and  
 XX encoding at least 8 amino acids of any of the 6888 amino acid sequences  
 XX fully defined in the specification. The probe is a single exon probe that  
 XX hybridises under high stringency conditions to a nucleic acid molecule  
 XX expressed in human cells or tissues. Also included are a spatially-  
 XX addressable set of single exon nucleic acid probes for measuring human  
 XX gene expression (comprising a plurality of single exon nucleic acid  
 XX probes cited above, where each of the plurality of probes is separately  
 XX and addressably isolatable or amplifiable from the plurality), a single  
 XX exon microarray for measuring human gene expression, a method of  
 XX measuring human gene expression, a vector comprising the single exon  
 XX probe cited above, an ORF-encoded peptide comprising at least 8  
 XX contiguous amino acids of any of the above-mentioned amino acid  
 XX sequences (optionally with conservative amino acid substitutions), an  
 XX isolated antibody that binds specifically to a peptide cited above,  
 XX methods of selling and/or licensing single exon probes or microarrays to  
 XX a customer desiring to measure gene expression, a method of providing  
 XX human gene expression data by subscription, and a computer-readable  
 XX storage medium which contains a database having a plurality of records  
 XX (each record including data on the expression of a single exon probe  
 XX cited above. The probe, methods and apparatus are useful in gene  
 XX expression analysis. The probes may be used as tools for surveying  
 XX tissues to detect the presence of expressed messages that contain their  
 XX specific exon, or in constructing genome-derived single exon microarrays.  
 XX In addition, the probes are used in identifying and characterising  
 XX alternative splicing events, in detecting and characterising gross  
 XX alterations in the genomic locus that includes their exon, in assessing  
 XX smaller genomic alterations, in priming the synthesis of nucleic acids,  
 XX or in expressing the ORF-encoded peptide. The present sequence is a human  
 XX single exon probe protein of the invention. Note: The sequence data for  
 XX this patent did not form part of the printed specification, but was  
 XX obtained in electronic format directly from USPTO at  
 XX seqdata.uspto.gov/sequence.html?DocID=20030194704  
 XX SQ Sequence 40 AA;  
 Query Match 4.0%; Score 13; DB 8; Length 40;  
 Best Local Similarity 100.0%; Pred. No. 7.8e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 183 FADDDNTYSLELF 195  
 Db 26 FADDDNTYSLELF 38  
 Search completed: May 9, 2006, 10:57:06  
 Job time : 58 secs

**This Page Blank (uspro)**